```
CC   from N. meningitidis strain EG329 is 1 of 10 NhhA polypeptide sequences
CC   (AAU06171-AAU06180) from 10 different N. meningitidis strains given in
CC   the present invention.
XX
SQ   Sequence   591 AA;

  Query Match             98.7%;  Score 2720;  DB 22;  Length 591;
  Best Local Similarity   99.1%;  Pred. No. 4.3e-157;
  Matches  535;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 NNEEQEEYLYLHPVQRTVAVLIVNSDKEGAGEKEKVEENSDWAVYFNEKGVLTAREITLK 60
              ||||||| ||| || ||||||||||||| |||||||||||||||||||||||||||||||
Db         52 NNEEQEEDLYLDPVLRTVAVLIVNSDKEGTGEKEKVEENSDWAVYFNEKGVLTAREITLK 111

Qy         61 AGDNLKIKQNGTNFTYSLKKDLTDLTSVGTEKLSFSAHGNKVNITSDTKGLNFAKETAGT 120
              ||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||||
Db        112 AGDNLKIKQNGTNFTYSLKKDLTDLTSVGTEKLSFSANGNKVNITSDTKGLNFAKETAGT 171

Qy        121 NGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        172 NGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPG 231

Qy        181 TTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        232 TTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTG 291

Qy        241 KDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        292 KDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFA 351

Qy        301 SGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSP 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        352 SGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSP 411

Qy        361 SKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGDALNVG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        412 SKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGDALNVG 471

Qy        421 SKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLV 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        472 SKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLV 531

Qy        481 QAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        532 QAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 591


RESULT 6
AAY57045
ID   AAY57045 standard; Protein; 591 AA.
XX
AC   AAY57045;
XX
DT   21-FEB-2000  (first entry)
XX
DE   BASB029 amino acid sequence from N. meningitidis strain H44/76.
XX
KW   BASB029; Nisseria meningitidis; surface fibril protein; HSF; diagnosis;
KW   infection; treatment; prevent; antibacterial drug.
XX
OS   Neisseria meningitidis.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 90
FT                   /note= "Encoded by AAT"
FT   Misc-difference 92
FT                   /note= "Encoded by GAT"
FT   Misc-difference 98
FT                   /note= "Encoded by AAC"
FT   Misc-difference 108
FT                   /note= "Encoded by AATC"
FT   Misc-difference 123
FT                   /note= "Encoded by ACA"
FT   Misc-difference 269
FT                   /note= "Encoded by AAA"
FT   Misc-difference 389
FT                   /note= "Encoded by CGT"
XX
PN   WO9958683-A2.
XX
PD   18-NOV-1999.
XX
PF   07-MAY-1999;   99WO-EP03255.
XX
PR   13-MAY-1998;   98GB-0010276.
XX
PA   (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PI   Ruelle J;
XX
DR   WPI; 2000-053103/04.
DR   N-PSDB; AAZ39865.
XX
PT   New polypeptide from neisseria meningitidis useful for diagnosis,
PT   treatment or prevention of bacterial infections in mammal  -
XX
PS   Claim 4; Fig 2; 74pp; English.
XX
CC   This is the Nisseria meningitidis BASB029 amino acid sequence from
CC   serogroup B strain H44/76. The BASB029 protein is homologous to the
CC   Haemophilus influenzae surface fibril (HSF) protein. The invention
CC   relates to BASB029 polynucleotide sequences (AAZ39864-Z39865) and
CC   polypeptide sequences (AAY57044-Y57045) and their immunogenic fragments.
CC   BASB029 polypeptides are useful in a method of diagnosing a Neisseria
CC   meningitidis infection in a mammal. Compositions containing BASB029
CC   polynucleotides and polypeptides are useful for generating an immune
CC   response in an animal. A therapeutic composition comprising an antibody
CC   directed against BASB029 is useful in treating humans with Neisseria
CC   meningitidis disease. The polynucleotide is useful in the diagnosis of
CC   the stage of infection, type of infection, susceptibility to an
CC   infection which results from increased or decreased expression of the
CC   polynucleotide, and for therapeutic or prophylactic purposes,
CC   particularly genetic immunisation. Antibodies against BASB029
CC   polynucleotides and polypeptides are also useful for treating infections
CC   particularly bacterial infections. The protein is useful in the
CC   screening and development of antibacterial drugs. Fused recombinant
CC   protein is useful for the stimulation of the immune system of an organism
CC   receiving the protein.
XX
SQ   Sequence   591 AA;

  Query Match             98.1%;  Score 2705;  DB 21;  Length 591;
  Best Local Similarity   98.3%;  Pred. No. 3.5e-156;
  Matches  531;  Conservative    6;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 NNEEQEEYLYLHPVQRTVAVLIVNSDKEGAGEKEKVEENSDWAVYFNEKGVLTAREITLK 60
              ||||||| ||| ||||||||||||||| |||||||:|:||||:||||||||||||||
Db         52 NNEEQEEDLYLDPVQRTVAVLIVNSDKEGTGEKEKVEEDSNWAVYFDEKGVLTAREITLK 111

Qy         61 AGDNLKIKQNGTNFTYSLKKDLTDLTSVGTEKLSFSAHGNKVNITSDTKGLNFAKETAGT 120
              ||||||||||:|||||||||||||||||||||||:||||||||||||||||||||||||||
Db        112 AGDNLKIKQNGSNFTYSLKKDLTDLTSVGTEKLSFSANGNKVNITSDTKGLNFAKETAGT 171

Qy        121 NGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        172 NGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPG 231

Qy        181 TTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTG 240
              ||||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||
Db        232 TTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKRTEVKIGAKTSVIKEKDGKLVTG 291

Qy        241 KDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        292 KDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFA 351

Qy        301 SGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSP 360
```

```
Db        352 SGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSP 411

Qy        361 SKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGDALNVG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        412 SKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGDALNVG 471

Qy        421 SKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLV 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        472 SKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLV 531

Qy        481 QAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        532 QAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 591


RESULT 7
AAY23737
ID   AAY23737 standard; Protein; 592 AA.
XX
AC   AAY23737;
XX
DT   08-SEP-1999  (first entry)
XX
DE   A surface protein of Neisseria meningitidis.
XX
KW   Surface protein; surface glycoprotein; infection; vaccine;
KW   immunoreactive peptide.
XX
OS   Neisseria meningitidis.
XX
PN   WO9931132-A1.
XX
PD   24-JUN-1999.
XX
PF   14-DEC-1998;   98WO-AU01031.
XX
PR   12-DEC-1997;   97GB-0026398.
XX
PA   (ISIS-) ISIS INNOVATION LTD.
PA   (UYQU ) UNIV QUEENSLAND.
XX
PI   Jennings MP,  Moxon ER,  Peak IRA;
XX
DR   WPI; 1999-418754/35.
DR   N-PSDB; AAX85788.
XX
PT   Neisseria meningitidis surface proteins useful for treating N.
PT   meningitidis infections
XX
PS   Claim 1; Page 86-87; 132pp; English.
XX
CC   The present sequence represents a surface protein of Neiserria
CC   meningitidis which is approximately 62 kDa. The N. meningitidis
CC   surface glycoproteins, nucleic acids, the primers and optionally
CC   a thermostable polymerase, or antibodies are useful in a kit for
CC   the detection or diagnosis of N. meningitidis infection in humans.
CC   The N. meningitidis surface glycoproteins can also be used to
CC   prevent or treat N. meningitidis infection in humans, especially
CC   in the form of vaccines. The proteins and antibodies can also
CC   be used to identify immunoreactive peptides.
XX
SQ   Sequence   592 AA;

  Query Match             97.8%;  Score 2695.5;  DB 20;  Length 592;
  Best Local Similarity   98.2%;  Pred. No. 1.3e-155;
  Matches  531;  Conservative    4;  Mismatches    5;  Indels    1;  Gaps    1;

Qy          1 NNEE-QEEYLYLHPVQRTVAVLIVNSDKEGAGEKEKVEENSDWAVYFNEKGVLTAREITL 59
              ||| ::: ||| ||||||||||||||| ||||||||||||||||||||||||||||||
Db         52 NNERPRKKDLYLDPVQRTVAVLIVNSDKEGTGEKEKVEENSDWAVYFNEKGVLTAREITL 111

Qy         60 KAGDNLKIKQNGTNFTYSLKKDLTDLTSVGTEKLSFSAHGNKVNITSDTKGLNFAKETAG 119
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        112 KAGDNLKIKQNGTNFTYSLKKDLTDLTSVGTEKLSFSANGNKVNITSDTKGLNFAKETAG 171

Qy        120 TNGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKP 179
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        172 TNGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKP 231

Qy        180 GTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVT 239
              |||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        232 GTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGVKTSVIKEKDGKLVT 291

Qy        240 GKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTF 299
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        292 GKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTF 351

Qy        300 ASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVS 359
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        352 ASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVS 411

Qy        360 PSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGDALNV 419
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        412 PSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGDALNV 471

Qy        420 GSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGL 479
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        472 GSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGL 531

Qy        480 VQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQ 539
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        532 VQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQ 591

Qy        540 W 540
              |
Db        592 W 592


RESULT 8
AAY23740
ID   AAY23740 standard; Protein; 594 AA.
XX
AC   AAY23740;
XX
DT   08-SEP-1999  (first entry)
XX
DE   A surface protein of Neisseria meningitidis.
XX
KW   Surface protein; surface glycoprotein; infection; vaccine;
KW   immunoreactive peptide.
XX
OS   Neisseria meningitidis.
XX
PN   WO9931132-A1.
XX
PD   24-JUN-1999.
XX
PF   14-DEC-1998;   98WO-AU01031.
XX
PR   12-DEC-1997;   97GB-0026398.
XX
PA   (ISIS-) ISIS INNOVATION LTD.
PA   (UYQU ) UNIV QUEENSLAND.
XX
PI   Jennings MP,  Moxon ER,  Peak IRA;
XX
DR   WPI; 1999-418754/35.
DR   N-PSDB; AAX85792.
XX
PT   Neisseria meningitidis surface proteins useful for treating N.
PT   meningitidis infections
XX
PS   Claim 1; Page 100-101; 132pp; English.
XX
CC   The present sequence represents a surface protein of Neiserria
```

```
CC   meningitidis which is approximately 62 kDa. The N. meningitidis
CC   surface glycoproteins, nucleic acids, the primers and optionally
CC   a thermostable polymerase, or antibodies are useful in a kit for
CC   the detection or diagnosis of N. meningitidis infection in humans.
CC   The N. meningitidis surface glycoproteins can also be used to
CC   prevent or treat N. meningitidis infection in humans, especially
CC   in the form of vaccines. The proteins and antibodies can also
CC   be used to identify immunoreactive peptides.
XX
SQ   Sequence   594 AA;

  Query Match             91.1%;  Score 2512;  DB 20;  Length 594;
  Best Local Similarity   92.4%;  Pred. No. 1.8e-144;
  Matches  501;  Conservative   13;  Mismatches   20;  Indels    8;  Gaps    3;

Qy          6 EEYLYLHPVQRTVAVLIVNSDKEGAGEKEKVEENSDWAVYFNEKGVLTAREITLKAGDNL 65
              :: ||| |||||  ||   ||||| |||| | |:|:| |||::|||||| ||||||||
Db         54 DDDLYLEPVQRTAVVLSFRSDKEGTGEKE-VTEDSNWGVYFDKKGVLTAGTITLKAGDNL 112

Qy         66 KIKQ------NGTNFTYSLKKDLTDLTSVGTEKLSFSAHGNKVNITSDTKGLNFAKETAG 119
              ||||      | ::|||||||||||||||||||||: ||||||||||||||:||
Db        113 KIKQNTNENTNASSFTYSLKKDLTDLTSVGTEKLSFSANSNKVNITSDTKGLNFAKKTAE 172

Qy        120 TNGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKP 179
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        173 TNGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKP 232

Qy        180 GTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVT 239
              ||||||||||||||||||||||||||||||||||||||:|||||||||||||||||||||
Db        233 GTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKRTEVKIGAKTSVIKEKDGKLVT 292

Qy        240 GKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTF 299
              ||||||| ||||:||||||||||||||||||||||||||||||||||||||||||||||
Db        293 GKDKGENDSSTDKGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTF 352

Qy        300 ASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVS 359
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        353 ASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVS 412

Qy        360 PSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGD-ALN 418
              ||||||||||||||||||||||||||||||||||||||||||||||||||| : |||
Db        413 PSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGALN 472

Qy        419 VGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAG 478
              ||||  |||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        473 VGSKDANKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNHIDNVDGNARAGIAQAIATAG 532

Qy        479 LVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGY 538
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        533 LVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGY 592

Qy        539 QW 540
              ||
Db        593 QW 594


RESULT 9
AAY57044
ID   AAY57044 standard; Protein; 594 AA.
XX
AC   AAY57044;
XX
DT   21-FEB-2000  (first entry)
XX
DE   BASB029 amino acid sequence from N. meningitidis strain ATCC13090.
XX
KW   BASB029; Nisseria meningitidis; surface fibril protein; HSF; diagnosis;
KW   infection; treatment; prevent; antibacterial drug.
XX
OS   Neisseria meningitidis.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 104
FT                   /note= "Encoded by AATC"
XX
PN   WO9958683-A2.
XX
PD   18-NOV-1999.
XX
PF   07-MAY-1999;   99WO-EP03255.
XX
PR   13-MAY-1998;   98GB-0010276.
XX
PA   (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PI   Ruelle J;
XX
DR   WPI; 2000-053103/04.
DR   N-PSDB; AAZ39864.
XX
PT   New polypeptide from neisseria meningitidis useful for diagnosis,
PT   treatment or prevention of bacterial infections in mammal -
XX
PS   Claim 4; Fig 2; 74pp; English.
XX
CC   This is the Nisseria meningitidis BASB029 amino acid sequence from
CC   serogroup B strain ATCC13090. The BASB029 protein is homologous to the
CC   Haemophilus influenzae surface fibril (HSF) protein. The invention
CC   relates to BASB029 polynucleotide sequences (AAZ39864-Z39865) and
CC   polypeptide sequences (AAY57044-Y57045) and their immunogenic fragments.
CC   BASB029 polypeptides are useful in a method of diagnosing a Neisseria
CC   meningitidis infection in a mammal. Compositions containing BASB029
CC   polynucleotides and polypeptides are useful for generating an immune
CC   response in an animal. A therapeutic composition comprising an antibody
CC   directed against BAsB029 is useful in treating humans with Neisseria
CC   meningitidis disease. The polynucleotide is useful in the diagnosis of
CC   the stage of infection, type of infection, susceptibility to an
CC   infection which results from increased or decreased expression of the
CC   polynucleotide, and for therapeutic or prophylactic purposes,
CC   particularly genetic immunisation. Antibodies against BASB029
CC   polynucleotides and polypeptides are also useful for treating infections
CC   particularly bacterial infections. The protein is useful in the
CC   screening and development of antibacterial drugs. Fused recombinant
CC   protein is useful for the stimulation of the immune system of an organism
CC   receiving the protein.
XX
SQ   Sequence   594 AA;

  Query Match             91.1%;  Score 2512;  DB 21;  Length 594;
  Best Local Similarity   92.4%;  Pred. No. 1.8e-144;
  Matches  501;  Conservative   13;  Mismatches   20;  Indels    8;  Gaps    3;

Qy          6 EEYLYLHPVQRTVAVLIVNSDKEGAGEKEKVEENSDWAVYFNEKGVLTAREITLKAGDNL 65
              :: ||| |||||  ||   ||||| |||| | |:|:| |||::|||||| ||||||||
Db         54 DDDLYLEPVQRTAVVLSFRSDKEGTGEKE-VTEDSNWGVYFDKKGVLTAGTITLKAGDNL 112

Qy         66 KIKQ------NGTNFTYSLKKDLTDLTSVGTEKLSFSAHGNKVNITSDTKGLNFAKETAG 119
              ||||      | ::|||||||||||||||||||||: ||||||||||||||:||
Db        113 KIKQNTNENTNASSFTYSLKKDLTDLTSVGTEKLSFSANSNKVNITSDTKGLNFAKKTAE 172

Qy        120 TNGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKP 179
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        173 TNGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKP 232

Qy        180 GTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVT 239
              ||||||||||||||||||||||||||||||||||||||:|||||||||||||||||||||
Db        233 GTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKRTEVKIGAKTSVIKEKDGKLVT 292

Qy        240 GKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTF 299
              ||||||| ||||:||||||||||||||||||||||||||||||||||||||||||||||
Db        293 GKDKGENDSSTDKGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTF 352

Qy        300 ASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVS 359
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        353 ASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVS 412
```

```
Qy        360 PSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGD-ALN 418
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||| : |||
Db        413 PSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGALN 472

Qy        419 VGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAG 478
              ||||  |||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        473 VGSKDANKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNHIDNVDGNARAGIAQAIATAG 532

Qy        479 LVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGY 538
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        533 LVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGY 592

Qy        539 QW 540
              ||
Db        593 QW 594
```

```
RESULT 10
AAU06174
ID   AAU06174 standard; Protein; 594 AA.
XX
AC   AAU06174;
XX
DT   24-OCT-2001  (first entry)
XX
DE   N. meningitidis EG327 surface antigen NhhA polypeptide sequence.
XX
KW   Surface antigen NhhA; meningococcal disease; meningitis vaccine.
XX
OS   Neisseria meningitidis strain EG327.
XX
FH   Key             Location/Qualifiers
FT   Region          1..50
FT                   /label= C1
FT                   /note= "Conserved region 1"
FT   Region          51..104
FT                   /label= V1
FT                   /note= "Variable region 1"
FT   Region          105..116
FT                   /label= C2
FT                   /note= "Conserved region 2"
FT   Region          117..126
FT                   /label= V2
FT                   /note= "Variable region 2"
FT   Region          127..190
FT                   /label= C3
FT                   /note= "Conserved region 3"
FT   Region          191..212
FT                   /label= V3
FT                   /note= "Variable region 3"
FT   Region          213..231
FT                   /label= C4
FT                   /note= "Conserved region 4"
FT   Region          232..238
FT                   /label= V4
FT                   /note= "Variable region 4"
FT   Region          239..594
FT                   /label= C5
FT                   /note= "Conserved region 5"
XX
PN   WO200155182-A1.
XX
PD   02-AUG-2001.
XX
PF   25-JAN-2001; 2001WO-AU00069.
XX
PR   25-JAN-2000; 2000US-0177917.
XX
PA   (UYQU ) UNIV QUEENSLAND.
XX
PI   Peak IRA,  Jennings MP;
XX
DR   WPI; 2001-488774/53.
DR   N-PSDB; AAS09164.
XX
PT   New NhhA surface antigen polypeptides and polynucleotides from
PT   Neisseria meningitidis, useful in producing vaccines for treating or
PT   preventing broad spectrum of Neisseria meningitidis -
XX
PS   Claim 9; Fig 1; 91pp; English.
XX
CC   The present invention relates to the isolation of novel Neisseria
CC   meningitidis mutant polypeptides of the surface antigen NhhA
CC   (AAU06182-AAU06186). The modified or mutant NhhA polypeptides are
CC   characterised by deletions of non-conserved amino acids, particularly
CC   the deletion of variable regions. The deletion mutants are useful in
CC   diagnostics, therapeutic and prophylactic vaccines against a broader
CC   spectrum of N. meningitidis, and in designing and/or screening of
CC   medicaments. The mutant proteins when used as a vaccine can effectively
CC   immunise against a broader spectrum of N. meningitidis strains than
CC   would be expected from a corresponding wild-type surface antigen.
CC   The present sequence representing the wild type surface antigen NhhA
CC   from N. meningitidis strain EG327 is 1 of 10 NhhA polypeptide sequences
CC   (AAU06171-AAU06180) from 10 different N. meningitidis strains given in
CC   the present invention.
XX
SQ   Sequence   594 AA;

  Query Match             91.1%;  Score 2512;  DB 22;  Length 594;
  Best Local Similarity   92.4%;  Pred. No. 1.8e-144;
  Matches  501;  Conservative   13;  Mismatches   20;  Indels    8;  Gaps    3;

Qy          6 EEYLYLHPVQRTVAVLIVNSDKEGAGEKEKVEENSDWAVYFNEKGVLTAREITLKAGDNL 65
              :: ||| ||||| ||   ||||| |||| | |:|:| |||::|||||| ||||||||
Db         54 DDDLYLEPVQRTAVVLSFRSDKEGTGEKE-VTEDSNWGVYFDKKGVLTAGTITLKAGDNL 112

Qy         66 KIKQ------NGTNFTYSLKKDLTDLTSVGTEKLSFSAHGNKVNITSDTKGLNFAKETAG 119
              ||||      | ::|||||||||||||||||||||: ||||||||||||||:||
Db        113 KIKQNTNENTNASSFTYSLKKDLTDLTSVGTEKLSFSANSNKVNITSDTKGLNFAKKTAE 172

Qy        120 TNGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKP 179
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        173 TNGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKP 232

Qy        180 GTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVT 239
              ||||||||||||||||||||||||||||||||||||:||||||||||||||||||||||
Db        233 GTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKRTEVKIGAKTSVIKEKDGKLVT 292

Qy        240 GKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTF 299
              ||||||| ||||:||||||||||||||||||||||||||||||||||||||||||||||
Db        293 GKDKGENDSSTDKGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTF 352

Qy        300 ASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVS 359
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        353 ASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVS 412

Qy        360 PSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGD-ALN 418
              |||||||||||||||||||||||||||||||||||||||||||||||||||||| : |||
Db        413 PSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGALN 472

Qy        419 VGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAG 478
              ||||  |||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        473 VGSKDANKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNHIDNVDGNARAGIAQAIATAG 532

Qy        479 LVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGY 538
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        533 LVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGY 592

Qy        539 QW 540
              ||
Db        593 QW 594
```

```
RESULT 11
AAY23739
ID   AAY23739 standard; Protein; 594 AA.
```

```
XX
AC   AAY23739;
XX
DT   08-SEP-1999  (first entry)
XX
DE   A surface protein of Neisseria meningitidis.
XX
KW   Surface protein; surface glycoprotein; infection; vaccine;
KW   immunoreactive peptide.
XX
OS   Neisseria meningitidis.
XX
PN   WO9931132-A1.
XX
PD   24-JUN-1999.
XX
PF   14-DEC-1998;   98WO-AU01031.
XX
PR   12-DEC-1997;   97GB-0026398.
XX
PA   (ISIS-) ISIS INNOVATION LTD.
PA   (UYQU ) UNIV QUEENSLAND.
XX
PI   Jennings MP,  Moxon ER,  Peak IRA;
XX
DR   WPI; 1999-418754/35.
DR   N-PSDB; AAX85791.
XX
PT   Neisseria meningitidis surface proteins useful for treating N.
PT   meningitidis infections
XX
PS   Claim 1; Page 95-97; 132pp; English.
XX
CC   The present sequence represents a surface protein of Neiserria
CC   meningitidis which is approximately 62 kDa. The N. meningitidis
CC   surface glycoproteins, nucleic acids, the primers and optionally
CC   a thermostable polymerase, or antibodies are useful in a kit for
CC   the detection or diagnosis of N. meningitidis infection in humans.
CC   The N. meningitidis surface glycoproteins can also be used to
CC   prevent or treat N. meningitidis infection in humans, especially
CC   in the form of vaccines. The proteins and antibodies can also
CC   be used to identify immunoreactive peptides.
XX
SQ   Sequence   594 AA;

  Query Match            90.4%;  Score 2493;  DB 20;  Length 594;
  Best Local Similarity  91.6%;  Pred. No. 2.6e-143;
  Matches  500;  Conservative    9;  Mismatches   29;  Indels    8;  Gaps    3;

Qy          2 NEEQEEYLYLHPVQRTVAVLIVNSDKEGAGEKEKVEENSDWAVYFNEKGVLTAREITLKA 61
              |   :: ||| |||||  ||   ||||| |||| |: |:|||||:|| || |  |||||
Db         50 NATDDDDLYLEPVQRTAVVLSFRSDKEGTGEKEGTED-SNWAVYFDEKRVLKAGAITLKA 108

Qy         62 GDNLKIKQ------NGTNFTYSLKKDLTDLTSVGTEKLSFSAHGNKVNITSDTKGLNFAK 115
              ||||||||      | ::|||||||||||||||| |||||| |:|||||||||||||||
Db        109 GDNLKIKQNTNENTNDSSFTYSLKKDLTDLTSVETEKLSFGANGNKVNITSDTKGLNFAK 168

Qy        116 ETAGTNGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIK 175
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        169 ETAGTNGDPTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIK 228

Qy        176 GVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDG 235
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        229 GVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDG 288

Qy        236 KLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGT 295
              |||||| | ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        289 KLVTGKGKDENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGT 348

Qy        296 NVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVIS 355
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        349 NVTFASGKGTTATVSKDDQGNITVKYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVIS 408

Qy        356 GNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGD 415
              |||||||||||||||||||||||||||||||||||||| |||||||||||||||||| :
Db        409 GNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMAPQFSSVSLGAGADAPTLSVDDE 468

Qy        416 -ALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAI 474
               |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        469 GALNVGSKDTNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAI 528

Qy        475 ATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASA 534
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        529 ATAGLVQAYLPGKSMMAIGGDTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASA 588

Qy        535 SVGYQW 540
              ||||||
Db        589 SVGYQW 594


RESULT 12
AAU06179
ID   AAU06179 standard; Protein; 594 AA.
XX
AC   AAU06179;
XX
DT   24-OCT-2001  (first entry)
XX
DE   N. meningitidis BZ198 surface antigen NhhA polypeptide sequence.
XX
KW   Surface antigen NhhA; meningococcal disease; meningitis vaccine.
XX
OS   Neisseria meningitidis strain BZ198.
XX
FH   Key             Location/Qualifiers
FT   Region          1..50
FT                   /label= C1
FT                   /note= "Conserved region 1"
FT   Region          51..104
FT                   /label= V1
FT                   /note= "Variable region 1"
FT   Region          105..116
FT                   /label= C2
FT                   /note= "Conserved region 2"
FT   Region          117..126
FT                   /label= V2
FT                   /note= "Variable region 2"
FT   Region          127..190
FT                   /label= C3
FT                   /note= "Conserved region 3"
FT   Region          191..212
FT                   /label= V3
FT                   /note= "Variable region 3"
FT   Region          213..231
FT                   /label= C4
FT                   /note= "Conserved region 4"
FT   Region          232..238
FT                   /label= V4
FT                   /note= "Variable region 4"
FT   Region          239..594
FT                   /label= C5
FT                   /note= "Conserved region 5"
XX
PN   WO200155182-A1.
XX
PD   02-AUG-2001.
XX
PF   25-JAN-2001; 2001WO-AU00069.
XX
PR   25-JAN-2000; 2000US-0177917.
XX
PA   (UYQU ) UNIV QUEENSLAND.
XX
PI   Peak IRA,  Jennings MP;
XX
DR   WPI; 2001-488774/53.
```

```
DR   N-PSDB; AAS09169.
XX
PT   New NhhA surface antigen polypeptides and polynucleotides from
PT   Neisseria meningitidis, useful in producing vaccines for treating or
PT   preventing broad spectrum of Neisseria meningitidis -
XX
PS   Claim 9; Fig 1; 91pp; English.
XX
CC   The present invention relates to the isolation of novel Neisseria
CC   meningitidis mutant polypeptides of the surface antigen NhhA
CC   (AAU06182-AAU06186). The modified or mutant NhhA polypeptides are
CC   characterised by deletions of non-conserved amino acids, particularly
CC   the deletion of variable regions. The deletion mutants are useful in
CC   diagnostics, therapeutic and prophylactic vaccines against a broader
CC   spectrum of N. meningitidis, and in designing and/or screening of
CC   medicaments. The mutant proteins when used as a vaccine can effectively
CC   immunise against a broader spectrum of N. meningitidis strains than
CC   would be expected from a corresponding wild-type surface antigen.
CC   The present sequence representing the wild type surface antigen NhhA
CC   from N. meningitidis strain BZ198 is 1 of 10 NhhA polypeptide sequences
CC   (AAU06171-AAU06180) from 10 different N. meningitidis strains given in
CC   the present invention.
XX
SQ   Sequence   594 AA;

  Query Match             90.4%;  Score 2493;  DB 22;  Length 594;
  Best Local Similarity   91.6%;  Pred. No. 2.6e-143;
  Matches  500;  Conservative    9;  Mismatches   29;  Indels    8;  Gaps    3;

Qy          2 NEEQEEYLYLHPVQRTVAVLIVNSDKEGAGEKEKVEENSDWAVYFNEKGVLTAREITLKA 61
              |   :: ||| ||||| ||   ||||| |||| |: |:|||||:|| || |  |||||
Db         50 NATDDDDLYLEPVQRTAVVLSFRSDKEGTGEKEGTED-SNWAVYFDEKRVLKAGAITLKA 108

Qy         62 GDNLKIKQ------NGTNFTYSLKKDLTDLTSVGTEKLSFSAHGNKVNITSDTKGLNFAK 115
              ||||||||      | ::||||||||||||||| |||||| |:||||||||||||||||
Db        109 GDNLKIKQNTNENTNDSSFTYSLKKDLTDLTSVETEKLSFGANGNKVNITSDTKGLNFAK 168

Qy        116 ETAGTNGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIK 175
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        169 ETAGTNGDPTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIK 228

Qy        176 GVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDG 235
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        229 GVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDG 288

Qy        236 KLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGT 295
              |||||| | ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        289 KLVTGKGKDENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGT 348

Qy        296 NVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVIS 355
              |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        349 NVTFASGKGTTATVSKDDQGNITVKYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVIS 408

Qy        356 GNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGD 415
              |||||||||||||||||||||||||||||||||||| |||||||||||||||||||| :
Db        409 GNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMAPQFSSVSLGAGADAPTLSVDDE 468

Qy        416 -ALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAI 474
               |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        469 GALNVGSKDTNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAI 528

Qy        475 ATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASA 534
              |||||||||||||||||| ||||||||||||||||||||||||\|||||||||||||||
Db        529 ATAGLVQAYLPGKSMMAIGGDTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASA 588

Qy        535 SVGYQW 540
              ||||||
Db        589 SVGYQW 594


RESULT 13
AAY23738
ID   AAY23738 standard; Protein; 598 AA.
XX
AC   AAY23738;
XX
DT   08-SEP-1999  (first entry)
XX
DE   A surface protein of Neisseria meningitidis.
XX
KW   Surface protein; surface glycoprotein; infection; vaccine;
KW   immunoreactive peptide.
XX
OS   Neisseria meningitidis.
XX
PN   WO9931132-A1.
XX
PD   24-JUN-1999.
XX
PF   14-DEC-1998;   98WO-AU01031.
XX
PR   12-DEC-1997;   97GB-0026398.
XX
PA   (ISIS-) ISIS INNOVATION LTD.
PA   (UYQU ) UNIV QUEENSLAND.
XX
PI   Jennings MP,  Moxon ER,  Peak IRA;
XX
DR   WPI; 1999-418754/35.
DR   N-PSDB; AAX85790.
XX
PT   Neisseria meningitidis surface proteins useful for treating N.
PT   meningitidis infections
XX
PS   Claim 1; Page 91-93; 132pp; English.
XX
CC   The present sequence represents a surface protein of Neiserria
CC   meningitidis which is approximately 62 kDa. The N. meningitidis
CC   surface glycoproteins, nucleic acids, the primers and optionally
CC   a thermostable polymerase, or antibodies are useful in a kit for
CC   the detection or diagnosis of N. meningitidis infection in humans.
CC   The N. meningitidis surface glycoproteins can also be used to
CC   prevent or treat N. meningitidis infection in humans, especially
CC   in the form of vaccines. The proteins and antibodies can also
CC   be used to identify immunoreactive peptides.
XX
SQ   Sequence   598 AA;

  Query Match             89.9%;  Score 2478;  DB 20;  Length 598;
  Best Local Similarity   90.2%;  Pred. No. 2.1e-142;
  Matches  496;  Conservative   11;  Mismatches   31;  Indels   12;  Gaps    3;

Qy          2 NEEQEEYLYLHPVQRTVAVLIVNSDKEGAGEKEKVEENSDWAVYFNEKGVLTAREITLKA 61
              |   :: ||| ||||| ||   ||||| |||| |: |:|||||:|| || |  |||||
Db         50 NATDDDDLYLEPVQRTAVVLSFRSDKEGTGEKEGTED-SNWAVYFDEKRVLKAGAITLKA 108

Qy         62 GDNLKIKQ----------NGTNFTYSLKKDLTDLTSVGTEKLSFSAHGNKVNITSDTKGL 111
              |||||||           | ::||||||||||||| |||||| |:||||||||||||
Db        109 GDNLKIKQNTNENTNENTNDSSFTYSLKKDLTDLTSVETEKLSFGANGNKVNITSDTKGL 168

Qy        112 NFAKETAGTNGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAG 171
              |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        169 NFAKETAGTNGDPTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAG 228

Qy        172 WNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIK 231
              ||||||||||||||||||||||||||||||||||||||||||||||:||||||||||||
Db        229 WNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKRTEVKIGAKTSVIK 288

Qy        232 EKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETV 291
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db        289 EKDGKLVTGKGKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETV 348

Qy        292 TSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSG 351
              |||| |||||| |||||||||||||| ||||||||||||||||||||||||||||||||
Db        349 TSGTKVTFASGNGTTATVSKDDQGNITVKYDVNVGDALNVNQLQNSGWNLDSKAVAGSSG 408
```

```
Qy        352 KVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLS 411
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        409 KVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLS 468

Qy        412 VDGD-ALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGI 470
              || : |||||||  ||||||||||||||||||||||||||||||||||||||||||||||
Db        469 VDDEGALNVGSKDANKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGI 528

Qy        471 AQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHF 530
              |||||||| |||||||||||||||||||||||||||| |||:||||||||||||||
Db        529 AQAIATAGLAQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDTGNWVIKGTASGNSRGHF 588

Qy        531 GASASVGYQW 540
              | ||||||||
Db        589 GTSASVGYQW 598
```

RESULT 14
AAY23742
ID AAY23742 standard; Protein; 598 AA.
XX
AC AAY23742;
XX
DT 08-SEP-1999 (first entry)
XX
DE A surface protein of Neisseria meningitidis.
XX
KW Surface protein; surface glycoprotein; infection; vaccine;
KW immunoreactive peptide.
XX
OS Neisseria meningitidis.
XX
PN WO9931132-A1.
XX
PD 24-JUN-1999.
XX
PF 14-DEC-1998; 98WO-AU01031.
XX
PR 12-DEC-1997; 97GB-0026398.
XX
PA (ISIS-) ISIS INNOVATION LTD.
PA (UYQU) UNIV QUEENSLAND.
XX
PI Jennings MP, Moxon ER, Peak IRA;
XX
DR WPI; 1999-418754/35.
DR N-PSDB; AAX85794.
XX
PT Neisseria meningitidis surface proteins useful for treating N.
PT meningitidis infections
XX
PS Claim 1; Page 108-110; 132pp; English.
XX
CC The present sequence represents a surface protein of Neiserria
CC meningitidis which is approximately 62 kDa. The N. meningitidis
CC surface glycoproteins, nucleic acids, the primers and optionally
CC a thermostable polymerase, or antibodies are useful in a kit for
CC the detection or diagnosis of N. meningitidis infection in humans.
CC The N. meningitidis surface glycoproteins can also be used to
CC prevent or treat N. meningitidis infection in humans, especially
CC in the form of vaccines. The proteins and antibodies can also
CC be used to identify immunoreactive peptides.
XX
SQ Sequence 598 AA;

Query Match 89.9%; Score 2478; DB 20; Length 598;
Best Local Similarity 90.4%; Pred. No. 2.1e-142;
Matches 497; Conservative 10; Mismatches 31; Indels 12; Gaps 3;

```
Qy          2 NEEQEEYLYLHPVQRTVAVLIVNSDKEGAGEKEKVEENSDWAVYFNEKGVLTAREITLKA 61
              |   :: ||| ||||| ||   ||||| |||| |: |:|||||:|| || |  |||||
Db         50 NATDDDDLYLEPVQRTAVVLSFRSDKEGTGEKEGTED-SNWAVYFDEKRVLKAGAITLKA 108

Qy         62 GDNLKIKQ----------NGTNFTYSLKKDLTDLTSVGTEKLSFSAHGNKVNITSDTKGL 111
              ||||||||          | ::|||||||||||||| |||||| |:||||||||||||
Db        109 GDNLKIKQNTNENTNENTNDSSFTYSLKKDLTDLTSVETEKLSFGANGNKVNITSDTKGL 168

Qy        112 NFAKETAGTNGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAG 171
              |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        169 NFAKETAGTNGDPTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAG 228

Qy        172 WNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIK 231
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        229 WNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIK 288

Qy        232 EKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETV 291
              |||||||||| | ||||||||||||||||||||||||||||||||||||||||||||||
Db        289 EKDGKLVTGKGKDENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETV 348

Qy        292 TSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSG 351
              |||| |||||| ||||||||||||||| ||||||||||||||||||||||||||||||
Db        349 TSGTKVTFASGNGTTATVSKDDQGNITVKYDVNVGDALNVNQLQNSGWNLDSKAVAGSSG 408

Qy        352 KVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLS 411
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        409 KVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLS 468

Qy        412 VDGD-ALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGI 470
              || : |||||||  ||||||||||||||||||||||||||||||||||||||||||||||
Db        469 VDDEGALNVGSKDANKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGI 528

Qy        471 AQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHF 530
              |||||||| |||||||||||||||||||||||||||| |||:||||||||||||||
Db        529 AQAIATAGLAQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDTGNWVIKGTASGNSRGHF 588

Qy        531 GASASVGYQW 540
              ||||||||||
Db        589 GASASVGYQW 598
```

RESULT 15
AAU06177
ID AAU06177 standard; Protein; 598 AA.
XX
AC AAU06177;
XX
DT 24-OCT-2001 (first entry)
XX
DE N. meningitidis H15 surface antigen NhhA polypeptide sequence.
XX
KW Surface antigen NhhA; meningococcal disease; meningitis vaccine.
XX
OS Neisseria meningitidis strain H15.
XX

FH	Key	Location/Qualifiers
FT	Region	1..50
FT		/label= C1
FT		/note= "Conserved region 1"
FT	Region	51..104
FT		/label= V1
FT		/note= "Variable region 1"
FT	Region	105..116
FT		/label= C2
FT		/note= "Conserved region 2"
FT	Region	117..130
FT		/label= V2
FT		/note= "Variable region 2"
FT	Region	131..194
FT		/label= C3
FT		/note= "Conserved region 3"
FT	Region	195..216
FT		/label= V3
FT		/note= "Variable region 3"
FT	Region	217..235
FT		/label= C4
FT		/note= "Conserved region 4"

```
FT   Region          236..242
FT                   /label= V4
FT                   /note= "Variable region 4"
FT   Region          243..598
FT                   /label= C5
FT                   /note= "Conserved region 5"
XX
PN   WO200155182-A1.
XX
PD   02-AUG-2001.
XX
PF   25-JAN-2001; 2001WO-AU00069.
XX
PR   25-JAN-2000; 2000US-0177917.
XX
PA   (UYQU ) UNIV QUEENSLAND.
XX
PI   Peak IRA, Jennings MP;
XX
DR   WPI; 2001-488774/53.
DR   N-PSDB; AAS09167.
XX
PT   New NhhA surface antigen polypeptides and polynucleotides from
PT   Neisseria meningitidis, useful in producing vaccines for treating or
PT   preventing broad spectrum of Neisseria meningitidis -
XX
PS   Claim 9; Fig 1; 91pp; English.
XX
CC   The present invention relates to the isolation of novel Neisseria
CC   meningitidis mutant polypeptides of the surface antigen NhhA
CC   (AAU06182-AAU06186). The modified or mutant NhhA polypeptides are
CC   characterised by deletions of non-conserved amino acids, particularly
CC   the deletion of variable regions. The deletion mutants are useful in
CC   diagnostics, therapeutic and prophylactic vaccines against a broader
CC   spectrum of N. meningitidis, and in designing and/or screening of
CC   medicaments. The mutant proteins when used as a vaccine can effectively
CC   immunise against a broader spectrum of N. meningitidis strains than
CC   would be expected from a corresponding wild-type surface antigen.
CC   The present sequence representing the wild type surface antigen NhhA
CC   from N. meningitidis strain H15 is 1 of 10 NhhA polypeptide sequences
CC   (AAU06171-AAU06180) from 10 different N. meningitidis strains given in
CC   the present invention.
XX
SQ   Sequence   598 AA;

  Query Match             89.9%;  Score 2478;  DB 22;  Length 598;
  Best Local Similarity   90.4%;  Pred. No. 2.1e-142;
  Matches  497;  Conservative   10;  Mismatches   31;  Indels   12;  Gaps    3;

Qy          2 NEEQEEYLYLHPVQRTVAVLIVNSDKEGAGEKEKVEENSDWAVYFNEKGVLTAREITLKA 61
              |   :: ||| |||||  ||   ||||| ||||  |: |:|||||:|| || |  |||||
Db         50 NATDDDDLYLEPVQRTAVVLSFRSDKEGTGEKEGTED-SNWAVYFDEKRVLKAGAITLKA 108

Qy         62 GDNLKIKQ----------NGTNFTYSLKKDLTDLTSVGTEKLSFSAHGNKVNITSDTKGL 111
              ||||||||          | ::|||||||||||||| |||||| |:||||||||||||
Db        109 GDNLKIKQNTNENTNENTNDSSFTYSLKKDLTDLTSVETEKLSFGANGNKVNITSDTKGL 168

Qy        112 NFAKETAGTNGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAG 171
              |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        169 NFAKETAGTNGDPTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAG 228

Qy        172 WNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIK 231
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        229 WNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIK 288

Qy        232 EKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETV 291
              |||||||||| | ||||||||||||||||||||||||||||||||||||||||||||||
Db        289 EKDGKLVTGKGKDENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETV 348

Qy        292 TSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSG 351
              |||| |||||| |||||||||||||| ||||||||||||||||||||||||||||||||
Db        349 TSGTKVTFASGNGTTATVSKDDQGNITVKYDVNVGDALNVNQLQNSGWNLDSKAVAGSSG 408

Qy        352 KVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLS 411
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        409 KVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLS 468

Qy        412 VDGD-ALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGI 470
              || : |||||||  ||||||||||||||||||||||||||||||||||||||||||||||
Db        469 VDDEGALNVGSKDANKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGI 528

Qy        471 AQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHF 530
              ||||||||| |||||||||||||||||||||||||||||||||| |||:||||||||||||
Db        529 AQAIATAGLAQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDTGNWVIKGTASGNSRGHF 588

Qy        531 GASASVGYQW 540
              ||||||||||
Db        589 GASASVGYQW 598


Search completed: October  6, 2003, 09:22:37
Job time : 47.3729 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: October 6, 2003, 09:14:45 ; Search time 14.4407 Seconds
(without alignments)
1582.188 Million cell updates/sec

Title: US-09-771-382-33
Perfect score: 2757
Sequence: 1 NNEEQEEYLYLHPVQRTVAV.....TASGNSRGHFGASASVGYQW 540

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2727	98.9	591	3	US-09-377-155-21	Sequence 21, Appl
2	2727	98.9	591	4	US-09-669-974-21	Sequence 21, Appl
3	2720	98.7	591	3	US-09-377-155-11	Sequence 11, Appl
4	2720	98.7	591	4	US-09-669-974-11	Sequence 11, Appl
5	2695.5	97.8	592	3	US-09-377-155-2	Sequence 2, Appli
6	2695.5	97.8	592	4	US-09-669-974-2	Sequence 2, Appli
7	2512	91.1	594	3	US-09-377-155-9	Sequence 9, Appli
8	2512	91.1	594	4	US-09-669-974-9	Sequence 9, Appli
9	2493	90.4	594	3	US-09-377-155-7	Sequence 7, Appli
10	2493	90.4	594	4	US-09-669-974-7	Sequence 7, Appli
11	2478	89.9	598	3	US-09-377-155-5	Sequence 5, Appli
12	2478	89.9	598	3	US-09-377-155-13	Sequence 13, Appl
13	2478	89.9	598	4	US-09-669-974-5	Sequence 5, Appli
14	2478	89.9	598	4	US-09-669-974-13	Sequence 13, Appl
15	2430.5	88.2	599	3	US-09-377-155-15	Sequence 15, Appl
16	2430.5	88.2	599	4	US-09-669-974-15	Sequence 15, Appl
17	2321	84.2	592	3	US-09-377-155-17	Sequence 17, Appl
18	2321	84.2	592	4	US-09-669-974-17	Sequence 17, Appl
19	2241.5	81.3	589	3	US-09-377-155-19	Sequence 19, Appl
20	2241.5	81.3	589	4	US-09-669-974-19	Sequence 19, Appl
21	1033.5	37.5	2411	4	US-09-268-347-36	Sequence 36, Appl
22	1031.5	37.4	2353	3	US-09-377-155-33	Sequence 33, Appl
23	1031.5	37.4	2353	3	US-08-913-942-4	Sequence 4, Appli
24	1031.5	37.4	2353	4	US-09-669-974-33	Sequence 33, Appl
25	1030.5	37.4	2354	4	US-09-268-347-47	Sequence 47, Appl
26	1019.5	37.0	607	1	US-08-409-995-6	Sequence 6, Appli
27	1019.5	37.0	607	3	US-08-685-467-6	Sequence 6, Appli
28	1019.5	37.0	607	3	US-08-913-942-6	Sequence 6, Appli
29	1019.5	37.0	1912	1	US-08-409-995-4	Sequence 4, Appli
30	1019.5	37.0	1912	3	US-08-685-467-4	Sequence 4, Appli
31	927	33.6	1094	4	US-09-268-347-32	Sequence 32, Appl
32	908	32.9	1098	1	US-08-409-995-2	Sequence 2, Appli
33	908	32.9	1098	3	US-08-685-467-2	Sequence 2, Appli
34	908	32.9	1098	3	US-09-377-155-32	Sequence 32, Appl
35	908	32.9	1098	3	US-08-913-942-2	Sequence 2, Appli
36	908	32.9	1098	4	US-09-669-974-32	Sequence 32, Appl
37	908	32.9	1098	4	US-09-268-347-44	Sequence 44, Appl
38	890	32.3	658	1	US-08-409-995-5	Sequence 5, Appli
39	890	32.3	658	3	US-08-685-467-5	Sequence 5, Appli
40	890	32.3	658	3	US-08-913-942-5	Sequence 5, Appli
41	809.5	29.4	679	3	US-08-913-942-15	Sequence 15, Appl
42	809.5	29.4	679	4	US-09-268-347-26	Sequence 26, Appl
43	598.5	21.7	1004	4	US-09-268-347-30	Sequence 30, Appl
44	593.5	21.5	1104	4	US-09-268-347-28	Sequence 28, Appl
45	593.5	21.5	1104	4	US-09-268-347-34	Sequence 34, Appl

ALIGNMENTS

RESULT 1
US-09-377-155-21
; Sequence 21, Application US/09377155
; Patent No. 6197312
; GENERAL INFORMATION:
; APPLICANT: PEAK, Ian Richard Anselm
; APPLICANT: JENNINGS, Michael Paul
; APPLICANT: MOXON, E. Richard
; TITLE OF INVENTION: NOVEL SURFACE ANTIGEN
; FILE REFERENCE: 065064/0128
; CURRENT APPLICATION NUMBER: US/09/377,155
; CURRENT FILING DATE: 1999-08-19
; PRIOR APPLICATION NUMBER: PCT/AU98/01031
; PRIOR FILING DATE: 1998-12-14
; PRIOR APPLICATION NUMBER: GB 9726398.2
; PRIOR FILING DATE: 1997-12-12
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 21
; LENGTH: 591
; TYPE: PRT
; ORGANISM: Neisseria meningitidis
US-09-377-155-21

Query Match 98.9%; Score 2727; DB 3; Length 591;
Best Local Similarity 99.3%; Pred. No. 3.2e-203;
Matches 536; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

```
Qy          1 NNEEQEEYLYLHPVQRTVAVLIVNSDKEGAGEKEKVEENSDWAVYFNEKGVLTAREITLK 60
              ||||||| ||| |||||||||||||||| ||||||||||||||||||||||||||||||
Db         52 NNEEQEEDLYLDPVQRTVAVLIVNSDKEGTGEKEKVEENSDWAVYFNEKGVLTAREITLK 111

Qy         61 AGDNLKIKQNGTNFTYSLKKDLTDLTSVGTEKLSFSAHGNKVNITSDTKGLNFAKETAGT 120
              ||||||||||||||||||||||||||||||||||||:|||||||||||||||||||||||
Db        112 AGDNLKIKQNGTNFTYSLKKDLTDLTSVGTEKLSFSANGNKVNITSDTKGLNFAKETAGT 171

Qy        121 NGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        172 NGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPG 231

Qy        181 TTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        232 TTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTG 291

Qy        241 KDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        292 KDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFA 351

Qy        301 SGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSP 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        352 SGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSP 411

Qy        361 SKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGDALNVG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        412 SKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGDALNVG 471

Qy        421 SKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLV 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        472 SKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLV 531

Qy        481 QAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        532 QAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 591


RESULT 2
US-09-669-974-21
; Sequence 21, Application US/09669974
; Patent No. 6333173
; GENERAL INFORMATION:
;  APPLICANT: PEAK, Ian Richard Anselm
;  APPLICANT:  JENNINGS, Michael Paul
;  APPLICANT:  MOXON, E. Richard
;  TITLE OF INVENTION: NOVEL SURFACE ANTIGEN
;  FILE REFERENCE: 065064/0128
;  CURRENT APPLICATION NUMBER: US/09/669,974
;  CURRENT FILING DATE:  2000-09-26
;  PRIOR APPLICATION NUMBER: US 09/377,155
;  PRIOR FILING DATE: 1999-08-19
;  PRIOR APPLICATION NUMBER: PCT/AU98/01031
;  PRIOR FILING DATE: 1998-12-14
;  PRIOR APPLICATION NUMBER: GB 9726398.2
;  PRIOR FILING DATE: 1997-12-12
;  NUMBER OF SEQ ID NOS: 33
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 21
;   LENGTH: 591
;   TYPE: PRT
;   ORGANISM: Neisseria meningitidis
US-09-669-974-21

  Query Match           98.9%;  Score 2727;  DB 4;  Length 591;
  Best Local Similarity   99.3%;  Pred. No. 3.2e-203;
  Matches  536;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 NNEEQEEYLYLHPVQRTVAVLIVNSDKEGAGEKEKVEENSDWAVYFNEKGVLTAREITLK 60
              ||||||| ||| ||||||||||||||| |||||||||||||||||||||||||||||||
Db         52 NNEEQEEDLYLDPVQRTVAVLIVNSDKEGTGEKEKVEENSDWAVYFNEKGVLTAREITLK 111

Qy         61 AGDNLKIKQNGTNFTYSLKKDLTDLTSVGTEKLSFSAHGNKVNITSDTKGLNFAKETAGT 120
              ||||||||||||||||||||||||||||||||||||:||||||||||||||||||||||
Db        112 AGDNLKIKQNGTNFTYSLKKDLTDLTSVGTEKLSFSANGNKVNITSDTKGLNFAKETAGT 171

Qy        121 NGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        172 NGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPG 231

Qy        181 TTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        232 TTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTG 291

Qy        241 KDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        292 KDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFA 351

Qy        301 SGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSP 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        352 SGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSP 411

Qy        361 SKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGDALNVG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        412 SKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGDALNVG 471

Qy        421 SKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLV 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        472 SKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLV 531

Qy        481 QAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        532 QAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 591


RESULT 3
US-09-377-155-11
; Sequence 11, Application US/09377155
; Patent No. 6197312
; GENERAL INFORMATION:
;  APPLICANT:  PEAK, Ian Richard Anselm
;  APPLICANT:  JENNINGS, Michael Paul
;  APPLICANT:  MOXON, E. Richard
;  TITLE OF INVENTION:  NOVEL SURFACE ANTIGEN
;  FILE REFERENCE:  065064/0128
;  CURRENT APPLICATION NUMBER: US/09/377,155
;  CURRENT FILING DATE:  1999-08-19
;  PRIOR APPLICATION NUMBER:  PCT/AU98/01031
;  PRIOR FILING DATE:  1998-12-14
;  PRIOR APPLICATION NUMBER:  GB 9726398.2
;  PRIOR FILING DATE:  1997-12-12
;  NUMBER OF SEQ ID NOS:  33
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 11
;   LENGTH: 591
;   TYPE: PRT
;   ORGANISM: Neisseria meningitidis
US-09-377-155-11

  Query Match           98.7%;  Score 2720;  DB 3;  Length 591;
  Best Local Similarity   99.1%;  Pred. No. 1.1e-202;
  Matches  535;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 NNEEQEEYLYLHPVQRTVAVLIVNSDKEGAGEKEKVEENSDWAVYFNEKGVLTAREITLK 60
              ||||||| ||| || |||||||||||||| |||||||||||||||||||||||||||||
Db         52 NNEEQEEDLYLDPVLRTVAVLIVNSDKEGTGEKEKVEENSDWAVYFNEKGVLTAREITLK 111

Qy         61 AGDNLKIKQNGTNFTYSLKKDLTDLTSVGTEKLSFSAHGNKVNITSDTKGLNFAKETAGT 120
              ||||||||||||||||||||||||||||||||||||:||||||||||||||||||||||
Db        112 AGDNLKIKQNGTNFTYSLKKDLTDLTSVGTEKLSFSANGNKVNITSDTKGLNFAKETAGT 171

Qy        121 NGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        172 NGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPG 231

Qy        181 TTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        232 TTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTG 291

Qy        241 KDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        292 KDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFA 351

Qy        301 SGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSP 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        352 SGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSP 411

Qy        361 SKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGDALNVG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        412 SKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGDALNVG 471

Qy        421 SKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLV 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        472 SKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLV 531

Qy        481 QAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        532 QAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 591
```

```
RESULT 4
US-09-669-974-11
; Sequence 11, Application US/09669974
; Patent No. 6333173
; GENERAL INFORMATION:
;  APPLICANT: PEAK, Ian Richard Anselm
;  APPLICANT:  JENNINGS, Michael Paul
;  APPLICANT:  MOXON, E. Richard
;  TITLE OF INVENTION: NOVEL SURFACE ANTIGEN
;  FILE REFERENCE: 065064/0128
;  CURRENT APPLICATION NUMBER: US/09/669,974
;  CURRENT FILING DATE:  2000-09-26
;  PRIOR APPLICATION NUMBER: US 09/377,155
;  PRIOR FILING DATE: 1999-08-19
;  PRIOR APPLICATION NUMBER: PCT/AU98/01031
;  PRIOR FILING DATE: 1998-12-14
;  PRIOR APPLICATION NUMBER: GB 9726398.2
;  PRIOR FILING DATE: 1997-12-12
;  NUMBER OF SEQ ID NOS: 33
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
;   LENGTH: 591
;   TYPE: PRT
;   ORGANISM: Neisseria meningitidis
US-09-669-974-11

  Query Match             98.7%;  Score 2720;  DB 4;  Length 591;
  Best Local Similarity   99.1%;  Pred. No. 1.1e-202;
  Matches  535;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 NNEEQEEYLYLHPVQRTVAVLIVNSDKEGAGEKEKVEENSDWAVYFNEKGVLTAREITLK 60
              ||||||| ||| || ||||||||||||| ||||||||||||||||||||||||||||||
Db         52 NNEEQEEDLYLDPVLRTVAVLIVNSDKEGTGEKEKVEENSDWAVYFNEKGVLTAREITLK 111

Qy         61 AGDNLKIKQNGTNFTYSLKKDLTDLTSVGTEKLSFSAHGNKVNITSDTKGLNFAKETAGT 120
              ||||||||||||||||||||||||||||||||||||:||||||||||||||||||||||
Db        112 AGDNLKIKQNGTNFTYSLKKDLTDLTSVGTEKLSFSANGNKVNITSDTKGLNFAKETAGT 171

Qy        121 NGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        172 NGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPG 231

Qy        181 TTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        232 TTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTG 291

Qy        241 KDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        292 KDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFA 351

Qy        301 SGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSP 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        352 SGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSP 411

Qy        361 SKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGDALNVG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        412 SKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGDALNVG 471

Qy        421 SKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLV 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        472 SKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLV 531

Qy        481 QAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        532 QAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 591


RESULT 5
US-09-377-155-2
; Sequence 2, Application US/09377155
; Patent No. 6197312
; GENERAL INFORMATION:
;  APPLICANT:  PEAK, Ian Richard Anselm
;  APPLICANT:  JENNINGS, Michael Paul
;  APPLICANT:  MOXON, E. Richard
;  TITLE OF INVENTION:  NOVEL SURFACE ANTIGEN
;  FILE REFERENCE:  065064/0128
;  CURRENT APPLICATION NUMBER: US/09/377,155
;  CURRENT FILING DATE:  1999-08-19
;  PRIOR APPLICATION NUMBER:  PCT/AU98/01031
;  PRIOR FILING DATE:  1998-12-14
;  PRIOR APPLICATION NUMBER:  GB 9726398.2
;  PRIOR FILING DATE:  1997-12-12
;  NUMBER OF SEQ ID NOS:  33
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 592
;   TYPE: PRT
;   ORGANISM: Neisseria meningitidis
US-09-377-155-2

  Query Match             97.8%;  Score 2695.5;  DB 3;  Length 592;
  Best Local Similarity   98.2%;  Pred. No. 8.9e-201;
  Matches  531;  Conservative    4;  Mismatches    5;  Indels    1;  Gaps    1;

Qy          1 NNEE-QEEYLYLHPVQRTVAVLIVNSDKEGAGEKEKVEENSDWAVYFNEKGVLTAREITL 59
              ||| ::: ||| |||||||||||||||| ||||||||||||||||||||||||||||||
Db         52 NNERPRKKDLYLDPVQRTVAVLIVNSDKEGTGEKEKVEENSDWAVYFNEKGVLTAREITL 111

Qy         60 KAGDNLKIKQNGTNFTYSLKKDLTDLTSVGTEKLSFSAHGNKVNITSDTKGLNFAKETAG 119
              ||||||||||||||||||||||||||||||||||||:||||||||||||||||||||||
Db        112 KAGDNLKIKQNGTNFTYSLKKDLTDLTSVGTEKLSFSANGNKVNITSDTKGLNFAKETAG 171

Qy        120 TNGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKP 179
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        172 TNGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKP 231

Qy        180 GTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVT 239
              ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db        232 GTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGVKTSVIKEKDGKLVT 291

Qy        240 GKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTF 299
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        292 GKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTF 351

Qy        300 ASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVS 359
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        352 ASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVS 411

Qy        360 PSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGDALNV 419
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        412 PSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGDALNV 471

Qy        420 GSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGL 479
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        472 GSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGL 531

Qy        480 VQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQ 539
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        532 VQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQ 591

Qy        540 W 540
              |
Db        592 W 592


RESULT 6
US-09-669-974-2
; Sequence 2, Application US/09669974
; Patent No. 6333173
; GENERAL INFORMATION:
;  APPLICANT: PEAK, Ian Richard Anselm
;  APPLICANT:  JENNINGS, Michael Paul
;  APPLICANT:  MOXON, E. Richard
;  TITLE OF INVENTION: NOVEL SURFACE ANTIGEN
```

```
;  FILE REFERENCE: 065064/0128
;  CURRENT APPLICATION NUMBER: US/09/669,974
;  CURRENT FILING DATE:  2000-09-26
;  PRIOR APPLICATION NUMBER: US 09/377,155
;  PRIOR FILING DATE: 1999-08-19
;  PRIOR APPLICATION NUMBER: PCT/AU98/01031
;  PRIOR FILING DATE: 1998-12-14
;  PRIOR APPLICATION NUMBER: GB 9726398.2
;  PRIOR FILING DATE: 1997-12-12
;  NUMBER OF SEQ ID NOS: 33
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 592
;   TYPE: PRT
;   ORGANISM: Neisseria meningitidis
US-09-669-974-2

  Query Match            97.8%;  Score 2695.5;  DB 4;  Length 592;
  Best Local Similarity  98.2%;  Pred. No. 8.9e-201;
  Matches  531;  Conservative    4;  Mismatches    5;  Indels    1;  Gaps    1;

Qy          1 NNEE-QEEYLYLHPVQRTVAVLIVNSDKEGAGEKEKVEENSDWAVYFNEKGVLTAREITL 59
              |||  ::: ||| |||||||||||||||| ||||||||||||||||||||||||||||||
Db         52 NNERPRKKDLYLDPVQRTVAVLIVNSDKEGTGEKEKVEENSDWAVYFNEKGVLTAREITL 111

Qy         60 KAGDNLKIKQNGTNFTYSLKKDLTDLTSVGTEKLSFSAHGNKVNITSDTKGLNFAKETAG 119
              ||||||||||||||||||||||||||||||||||||:||||||||||||||||||||||
Db        112 KAGDNLKIKQNGTNFTYSLKKDLTDLTSVGTEKLSFSANGNKVNITSDTKGLNFAKETAG 171

Qy        120 TNGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKP 179
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        172 TNGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKP 231

Qy        180 GTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVT 239
              ||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        232 GTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGVKTSVIKEKDGKLVT 291

Qy        240 GKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTF 299
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        292 GKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTF 351

Qy        300 ASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVS 359
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        352 ASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVS 411

Qy        360 PSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGDALNV 419
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        412 PSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGDALNV 471

Qy        420 GSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGL 479
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        472 GSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGL 531

Qy        480 VQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQ 539
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        532 VQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQ 591

Qy        540 W 540
              |
Db        592 W 592


RESULT 7
US-09-377-155-9
; Sequence 9, Application US/09377155
; Patent No. 6197312
; GENERAL INFORMATION:
;  APPLICANT:  PEAK, Ian Richard Anselm
;  APPLICANT:  JENNINGS, Michael Paul
;  APPLICANT:  MOXON, E. Richard
;  TITLE OF INVENTION:  NOVEL SURFACE ANTIGEN
;  FILE REFERENCE:  065064/0128
;  CURRENT APPLICATION NUMBER: US/09/377,155
;  CURRENT FILING DATE:  1999-08-19
;  PRIOR APPLICATION NUMBER:  PCT/AU98/01031
;  PRIOR FILING DATE:  1998-12-14
;  PRIOR APPLICATION NUMBER:  GB 9726398.2
;  PRIOR FILING DATE:  1997-12-12
;  NUMBER OF SEQ ID NOS:  33
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 9
;   LENGTH: 594
;   TYPE: PRT
;   ORGANISM: Neisseria meningitidis
US-09-377-155-9

  Query Match            91.1%;  Score 2512;  DB 3;  Length 594;
  Best Local Similarity  92.4%;  Pred. No. 1.5e-186;
  Matches  501;  Conservative   13;  Mismatches   20;  Indels    8;  Gaps    3;

Qy          6 EEYLYLHPVQRTVAVLIVNSDKEGAGEKEKVEENSDWAVYFNEKGVLTAREITLKAGDNL 65
              :: ||| |||||  ||   ||||| |||| | |:|:| |||::|||||  ||||||||
Db         54 DDDLYLEPVQRTAVVLSFRSDKEGTGEKE-VTEDSNWGVYFDKKGVLTAGTITLKAGDNL 112

Qy         66 KIKQ------NGTNFTYSLKKDLTDLTSVGTEKLSFSAHGNKVNITSDTKGLNFAKETAG 119
              ||||      | ::|||||||||||||||||||||: |||||||||||||||:||
Db        113 KIKQNTNENTNASSFTYSLKKDLTDLTSVGTEKLSFSANSNKVNITSDTKGLNFAKKTAE 172

Qy        120 TNGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKP 179
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        173 TNGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKP 232

Qy        180 GTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVT 239
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        233 GTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKRTEVKIGAKTSVIKEKDGKLVT 292

Qy        240 GKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTF 299
              ||||||| ||||:|||||||||||||||||||||||||||||||||||||||||||||||
Db        293 GKDKGENDSSTDKGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTF 352

Qy        300 ASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVS 359
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        353 ASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVS 412

Qy        360 PSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGD-ALN 418
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||| : |||
Db        413 PSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGALN 472

Qy        419 VGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAG 478
              ||||  |||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        473 VGSKDANKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNHIDNVDGNARAGIAQAIATAG 532

Qy        479 LVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGY 538
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        533 LVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGY 592

Qy        539 QW 540
              ||
Db        593 QW 594


RESULT 8
US-09-669-974-9
; Sequence 9, Application US/09669974
; Patent No. 6333173
; GENERAL INFORMATION:
;  APPLICANT: PEAK, Ian Richard Anselm
;  APPLICANT:  JENNINGS, Michael Paul
;  APPLICANT:  MOXON, E. Richard
;  TITLE OF INVENTION: NOVEL SURFACE ANTIGEN
;  FILE REFERENCE: 065064/0128
;  CURRENT APPLICATION NUMBER: US/09/669,974
;  CURRENT FILING DATE:  2000-09-26
;  PRIOR APPLICATION NUMBER: US 09/377,155
;  PRIOR FILING DATE: 1999-08-19
;  PRIOR APPLICATION NUMBER: PCT/AU98/01031
```

```
;  PRIOR FILING DATE: 1998-12-14
;  PRIOR APPLICATION NUMBER: GB 9726398.2
;  PRIOR FILING DATE: 1997-12-12
;  NUMBER OF SEQ ID NOS: 33
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
;   LENGTH: 594
;   TYPE: PRT
;   ORGANISM: Neisseria meningitidis
US-09-669-974-9

  Query Match             91.1%;  Score 2512;  DB 4;  Length 594;
  Best Local Similarity   92.4%;  Pred. No. 1.5e-186;
  Matches  501;  Conservative   13;  Mismatches   20;  Indels    8;  Gaps    3;

Qy          6 EEYLYLHPVQRTVAVLIVNSDKEGAGEKEKVEENSDWAVYFNEKGVLTAREITLKAGDNL 65
              :: ||| ||||| ||   ||||| |||| | |:|:| |||::|||||| ||||||||
Db         54 DDDLYLEPVQRTAVVLSFRSDKEGTGEKE-VTEDSNWGVYFDKKGVLTAGTITLKAGDNL 112

Qy         66 KIKQ------NGTNFTYSLKKDLTDLTSVGTEKLSFSAHGNKVNITSDTKGLNFAKETAG 119
              ||||      | ::|||||||||||||||||||||||: ||||||||||||||:||
Db        113 KIKQNTNENTNASSFTYSLKKDLTDLTSVGTEKLSFSANSNKVNITSDTKGLNFAKKTAE 172

Qy        120 TNGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKP 179
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        173 TNGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKP 232

Qy        180 GTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVT 239
              ||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||
Db        233 GTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKRTEVKIGAKTSVIKEKDGKLVT 292

Qy        240 GKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTF 299
              ||||||| ||||:|||||||||||||||||||||||||||||||||||||||||||||||
Db        293 GKDKGENDSSTDKGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTF 352

Qy        300 ASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVS 359
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        353 ASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVS 412

Qy        360 PSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGD-ALN 418
              |||||||||||||||||||||||||||||||||||||||||||||||||||| : |||
Db        413 PSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGALN 472

Qy        419 VGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAG 478
              ||||  |||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        473 VGSKDANKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNHIDNVDGNARAGIAQAIATAG 532

Qy        479 LVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGY 538
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        533 LVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGY 592

Qy        539 QW 540
              ||
Db        593 QW 594


RESULT 9
US-09-377-155-7
; Sequence 7, Application US/09377155
; Patent No. 6197312
; GENERAL INFORMATION:
;  APPLICANT:  PEAK, Ian Richard Anselm
;  APPLICANT:  JENNINGS, Michael Paul
;  APPLICANT:  MOXON, E. Richard
;  TITLE OF INVENTION:  NOVEL SURFACE ANTIGEN
;  FILE REFERENCE:  065064/0128
;  CURRENT APPLICATION NUMBER: US/09/377,155
;  CURRENT FILING DATE:  1999-08-19
;  PRIOR APPLICATION NUMBER:  PCT/AU98/01031
;  PRIOR FILING DATE:  1998-12-14
;  PRIOR APPLICATION NUMBER:  GB 9726398.2
;  PRIOR FILING DATE:  1997-12-12
;  NUMBER OF SEQ ID NOS:  33
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 7
;   LENGTH: 594
;   TYPE: PRT
;   ORGANISM: Neisseria meningitidis
US-09-377-155-7

  Query Match             90.4%;  Score 2493;  DB 3;  Length 594;
  Best Local Similarity   91.6%;  Pred. No. 4.4e-185;
  Matches  500;  Conservative    9;  Mismatches   29;  Indels    8;  Gaps    3;

Qy          2 NEEQEEYLYLHPVQRTVAVLIVNSDKEGAGEKEKVEENSDWAVYFNEKGVLTAREITLKA 61
              |   :: ||| ||||| ||   ||||| |||| |: |:|||||:|| || |  |||||
Db         50 NATDDDDLYLEPVQRTAVVLSFRSDKEGTGEKEGTED-SNWAVYFDEKRVLKAGAITLKA 108

Qy         62 GDNLKIKQ------NGTNFTYSLKKDLTDLTSVGTEKLSFSAHGNKVNITSDTKGLNFAK 115
              ||||||||      | ::|||||||||||||||| |||||| |:|||||||||||||||
Db        109 GDNLKIKQNTNENTNDSSFTYSLKKDLTDLTSVETEKLSFGANGNKVNITSDTKGLNFAK 168

Qy        116 ETAGTNGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIK 175
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        169 ETAGTNGDPTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIK 228

Qy        176 GVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDG 235
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        229 GVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDG 288

Qy        236 KLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGT 295
              |||||| | ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        289 KLVTGKGKDENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGT 348

Qy        296 NVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVIS 355
              |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        349 NVTFASGKGTTATVSKDDQGNITVKYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVIS 408

Qy        356 GNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGD 415
              |||||||||||||||||||||||||||||||||||||||||||||| |||||||||||| :
Db        409 GNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMAPQFSSVSLGAGADAPTLSVDDE 468

Qy        416 -ALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAI 474
               |||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        469 GALNVGSKDTNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAI 528

Qy        475 ATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASA 534
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        529 ATAGLVQAYLPGKSMMAIGGDTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASA 588

Qy        535 SVGYQW 540
              ||||||
Db        589 SVGYQW 594


RESULT 10
US-09-669-974-7
; Sequence 7, Application US/09669974
; Patent No. 6333173
; GENERAL INFORMATION:
;  APPLICANT: PEAK, Ian Richard Anselm
;  APPLICANT:  JENNINGS, Michael Paul
;  APPLICANT:  MOXON, E. Richard
;  TITLE OF INVENTION: NOVEL SURFACE ANTIGEN
;  FILE REFERENCE: 065064/0128
;  CURRENT APPLICATION NUMBER: US/09/669,974
;  CURRENT FILING DATE:  2000-09-26
;  PRIOR APPLICATION NUMBER: US 09/377,155
;  PRIOR FILING DATE: 1999-08-19
;  PRIOR APPLICATION NUMBER: PCT/AU98/01031
;  PRIOR FILING DATE: 1998-12-14
;  PRIOR APPLICATION NUMBER: GB 9726398.2
;  PRIOR FILING DATE: 1997-12-12
;  NUMBER OF SEQ ID NOS: 33
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
```

```
;   LENGTH: 594
;   TYPE: PRT
;   ORGANISM: Neisseria meningitidis
US-09-669-974-7

  Query Match           90.4%;  Score 2493;  DB 4;  Length 594;
  Best Local Similarity   91.6%;  Pred. No. 4.4e-185;
  Matches  500;  Conservative    9;  Mismatches   29;  Indels    8;  Gaps    3;

Qy          2 NEEQEEYLYLHPVQRTVAVLIVNSDKEGAGEKEKVEENSDWAVYFNEKGVLTAREITLKA 61
              |   :: ||| |||||  ||   ||||| |||| |: |:|||||:|| || |  |||||
Db         50 NATDDDDLYLEPVQRTAVVLSFRSDKEGTGEKEGTED-SNWAVYFDEKRVLKAGAITLKA 108

Qy         62 GDNLKIKQ------NGTNFTYSLKKDLTDLTSVGTEKLSFSAHGNKVNITSDTKGLNFAK 115
              ||||||||      | ::||||||||||||||| |||||| |:||||||||||||||||
Db        109 GDNLKIKQNTNENTNDSSFTYSLKKDLTDLTSVETEKLSFGANGNKVNITSDTKGLNFAK 168

Qy        116 ETAGTNGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIK 175
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        169 ETAGTNGDPTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIK 228

Qy        176 GVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDG 235
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        229 GVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDG 288

Qy        236 KLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGT 295
              |||||| | ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        289 KLVTGKGKDENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGT 348

Qy        296 NVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVIS 355
              |||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db        349 NVTFASGKGTTATVSKDDQGNITVKYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVIS 408

Qy        356 GNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGD 415
              |||||||||||||||||||||||||||||||||| |||||||||||||||||||||| :
Db        409 GNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMAPQFSSVSLGAGADAPTLSVDDE 468

Qy        416 -ALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAI 474
               |||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        469 GALNVGSKDTNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAI 528

Qy        475 ATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASA 534
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        529 ATAGLVQAYLPGKSMMAIGGDTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASA 588

Qy        535 SVGYQW 540
              ||||||
Db        589 SVGYQW 594


RESULT 11
US-09-377-155-5
; Sequence 5, Application US/09377155
; Patent No. 6197312
; GENERAL INFORMATION:
;  APPLICANT:  PEAK, Ian Richard Anselm
;  APPLICANT:  JENNINGS, Michael Paul
;  APPLICANT:  MOXON, E. Richard
;  TITLE OF INVENTION:  NOVEL SURFACE ANTIGEN
;  FILE REFERENCE:  065064/0128
;  CURRENT APPLICATION NUMBER: US/09/377,155
;  CURRENT FILING DATE:  1999-08-19
;  PRIOR APPLICATION NUMBER:  PCT/AU98/01031
;  PRIOR FILING DATE:  1998-12-14
;  PRIOR APPLICATION NUMBER:  GB 9726398.2
;  PRIOR FILING DATE:  1997-12-12
;  NUMBER OF SEQ ID NOS:  33
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 5
;   LENGTH: 598
;   TYPE: PRT
;   ORGANISM: Neisseria meningitidis
US-09-377-155-5

  Query Match           89.9%;  Score 2478;  DB 3;  Length 598;
  Best Local Similarity   90.2%;  Pred. No. 6.5e-184;
  Matches  496;  Conservative   11;  Mismatches   31;  Indels   12;  Gaps    3;

Qy          2 NEEQEEYLYLHPVQRTVAVLIVNSDKEGAGEKEKVEENSDWAVYFNEKGVLTAREITLKA 61
              |   :: ||| |||||  ||   ||||| |||| |: |:|||||:|| || |  |||||
Db         50 NATDDDDLYLEPVQRTAVVLSFRSDKEGTGEKEGTED-SNWAVYFDEKRVLKAGAITLKA 108

Qy         62 GDNLKIKQ----------NGTNFTYSLKKDLTDLTSVGTEKLSFSAHGNKVNITSDTKGL 111
              ||||||||          | ::||||||||||||||| |||||| |:||||||||||||
Db        109 GDNLKIKQNTNENTNENTNDSSFTYSLKKDLTDLTSVETEKLSFGANGNKVNITSDTKGL 168

Qy        112 NFAKETAGTNGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAG 171
              |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db        169 NFAKETAGTNGDPTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAG 228

Qy        172 WNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIK 231
              ||||||||||||||||||||||||||||||||||||||||||||||:||||||||||||||
Db        229 WNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKRTEVKIGAKTSVIK 288

Qy        232 EKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETV 291
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        289 EKDGKLVTGKGKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETV 348

Qy        292 TSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSG 351
              |||| |||||| |||||||||||||| ||||||||||||||||||||||||||||||||
Db        349 TSGTKVTFASGNGTTATVSKDDQGNITVKYDVNVGDALNVNQLQNSGWNLDSKAVAGSSG 408

Qy        352 KVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLS 411
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        409 KVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLS 468

Qy        412 VDGD-ALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGI 470
              || : |||||||  ||||||||||||||||||||||||||||||||||||||||||||||||
Db        469 VDDEGALNVGSKDANKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGI 528

Qy        471 AQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHF 530
              ||||||||| |||||||||||||||||||||||||||||| |||:||||||||||||||||
Db        529 AQAIATAGLAQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDTGNWVIKGTASGNSRGHF 588

Qy        531 GASASVGYQW 540
              | ||||||||
Db        589 GTSASVGYQW 598


RESULT 12
US-09-377-155-13
; Sequence 13, Application US/09377155
; Patent No. 6197312
; GENERAL INFORMATION:
;  APPLICANT:  PEAK, Ian Richard Anselm
;  APPLICANT:  JENNINGS, Michael Paul
;  APPLICANT:  MOXON, E. Richard
;  TITLE OF INVENTION:  NOVEL SURFACE ANTIGEN
;  FILE REFERENCE:  065064/0128
;  CURRENT APPLICATION NUMBER: US/09/377,155
;  CURRENT FILING DATE:  1999-08-19
;  PRIOR APPLICATION NUMBER:  PCT/AU98/01031
;  PRIOR FILING DATE:  1998-12-14
;  PRIOR APPLICATION NUMBER:  GB 9726398.2
;  PRIOR FILING DATE:  1997-12-12
;  NUMBER OF SEQ ID NOS:  33
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 13
;   LENGTH: 598
;   TYPE: PRT
;   ORGANISM: Neisseria meningitidis
US-09-377-155-13

  Query Match           89.9%;  Score 2478;  DB 3;  Length 598;
  Best Local Similarity   90.4%;  Pred. No. 6.5e-184;
  Matches  497;  Conservative   10;  Mismatches   31;  Indels   12;  Gaps    3;
```

```
Qy          2 NEEQEEYLYLHPVQRTVAVLIVNSDKEGAGEKEKVEENSDWAVYFNEKGVLTAREITLKA 61
              |   :: ||| |||||  ||   ||||| ||||  |: |:|||||:|| || |  |||||
Db         50 NATDDDDLYLEPVQRTAVVLSFRSDKEGTGEKEGTED-SNWAVYFDEKRVLKAGAITLKA 108

Qy         62 GDNLKIKQ----------NGTNFTYSLKKDLTDLTSVGTEKLSFSAHGNKVNITSDTKGL 111
              ||||||||          | ::|||||||||||||| |||||| |:||||||||||||
Db        109 GDNLKIKQNTNENTNENTNDSSFTYSLKKDLTDLTSVETEKLSFGANGNKVNITSDTKGL 168

Qy        112 NFAKETAGTNGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAG 171
              ||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db        169 NFAKETAGTNGDPTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAG 228

Qy        172 WNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIK 231
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        229 WNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIK 288

Qy        232 EKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETV 291
              |||||||||| | ||||||||||||||||||||||||||||||||||||||||||||||
Db        289 EKDGKLVTGKGKDENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETV 348

Qy        292 TSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSG 351
              |||| |||||| ||||||||||||||| ||||||||||||||||||||||||||||||||
Db        349 TSGTKVTFASGNGTTATVSKDDQGNITVKYDVNVGDALNVNQLQNSGWNLDSKAVAGSSG 408

Qy        352 KVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLS 411
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        409 KVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLS 468

Qy        412 VDGD-ALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGI 470
              || : |||||||  ||||||||||||||||||||||||||||||||||||||||||||||
Db        469 VDDEGALNVGSKDANKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGI 528

Qy        471 AQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHF 530
              |||||||| |||||||||||||||||||||||||||||||| |||:||||||||||||||
Db        529 AQAIATAGLAQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDTGNWVIKGTASGNSRGHF 588

Qy        531 GASASVGYQW 540
              ||||||||||
Db        589 GASASVGYQW 598
```

```
RESULT 13
US-09-669-974-5
; Sequence 5, Application US/09669974
; Patent No. 6333173
; GENERAL INFORMATION:
;  APPLICANT: PEAK, Ian Richard Anselm
;  APPLICANT:  JENNINGS, Michael Paul
;  APPLICANT:  MOXON, E. Richard
;  TITLE OF INVENTION: NOVEL SURFACE ANTIGEN
;  FILE REFERENCE: 065064/0128
;  CURRENT APPLICATION NUMBER: US/09/669,974
;  CURRENT FILING DATE:  2000-09-26
;  PRIOR APPLICATION NUMBER: US 09/377,155
;  PRIOR FILING DATE: 1999-08-19
;  PRIOR APPLICATION NUMBER: PCT/AU98/01031
;  PRIOR FILING DATE: 1998-12-14
;  PRIOR APPLICATION NUMBER: GB 9726398.2
;  PRIOR FILING DATE: 1997-12-12
;  NUMBER OF SEQ ID NOS: 33
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
;   LENGTH: 598
;   TYPE: PRT
;   ORGANISM: Neisseria meningitidis
US-09-669-974-5

  Query Match            89.9%;  Score 2478;  DB 4;  Length 598;
  Best Local Similarity   90.2%;  Pred. No. 6.5e-184;
  Matches  496;  Conservative   11;  Mismatches   31;  Indels   12;  Gaps    3;

Qy          2 NEEQEEYLYLHPVQRTVAVLIVNSDKEGAGEKEKVEENSDWAVYFNEKGVLTAREITLKA 61
```

```
              |   :: ||| |||||  ||   ||||| ||||  |: |:|||||:|| || |  |||||
Db         50 NATDDDDLYLEPVQRTAVVLSFRSDKEGTGEKEGTED-SNWAVYFDEKRVLKAGAITLKA 108

Qy         62 GDNLKIKQ----------NGTNFTYSLKKDLTDLTSVGTEKLSFSAHGNKVNITSDTKGL 111
              ||||||||          | ::|||||||||||||| |||||| |:||||||||||||
Db        109 GDNLKIKQNTNENTNENTNDSSFTYSLKKDLTDLTSVETEKLSFGANGNKVNITSDTKGL 168

Qy        112 NFAKETAGTNGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAG 171
              ||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db        169 NFAKETAGTNGDPTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAG 228

Qy        172 WNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIK 231
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        229 WNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKRTEVKIGAKTSVIK 288

Qy        232 EKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETV 291
              |||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||
Db        289 EKDGKLVTGKGKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETV 348

Qy        292 TSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSG 351
              |||| |||||| ||||||||||||||| ||||||||||||||||||||||||||||||||
Db        349 TSGTKVTFASGNGTTATVSKDDQGNITVKYDVNVGDALNVNQLQNSGWNLDSKAVAGSSG 408

Qy        352 KVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLS 411
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        409 KVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLS 468

Qy        412 VDGD-ALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGI 470
              || : |||||||  ||||||||||||||||||||||||||||||||||||||||||||||
Db        469 VDDEGALNVGSKDANKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGI 528

Qy        471 AQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHF 530
              |||||||| |||||||||||||||||||||||||||||||| |||:||||||||||||||
Db        529 AQAIATAGLAQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDTGNWVIKGTASGNSRGHF 588

Qy        531 GASASVGYQW 540
              ||||||||||
Db        589 GTSASVGYQW 598
```

```
RESULT 14
US-09-669-974-13
; Sequence 13, Application US/09669974
; Patent No. 6333173
; GENERAL INFORMATION:
;  APPLICANT: PEAK, Ian Richard Anselm
;  APPLICANT:  JENNINGS, Michael Paul
;  APPLICANT:  MOXON, E. Richard
;  TITLE OF INVENTION: NOVEL SURFACE ANTIGEN
;  FILE REFERENCE: 065064/0128
;  CURRENT APPLICATION NUMBER: US/09/669,974
;  CURRENT FILING DATE:  2000-09-26
;  PRIOR APPLICATION NUMBER: US 09/377,155
;  PRIOR FILING DATE: 1999-08-19
;  PRIOR APPLICATION NUMBER: PCT/AU98/01031
;  PRIOR FILING DATE: 1998-12-14
;  PRIOR APPLICATION NUMBER: GB 9726398.2
;  PRIOR FILING DATE: 1997-12-12
;  NUMBER OF SEQ ID NOS: 33
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
;   LENGTH: 598
;   TYPE: PRT
;   ORGANISM: Neisseria meningitidis
US-09-669-974-13

  Query Match            89.9%;  Score 2478;  DB 4;  Length 598;
  Best Local Similarity   90.4%;  Pred. No. 6.5e-184;
  Matches  497;  Conservative   10;  Mismatches   31;  Indels   12;  Gaps    3;

Qy          2 NEEQEEYLYLHPVQRTVAVLIVNSDKEGAGEKEKVEENSDWAVYFNEKGVLTAREITLKA 61
              |   :: ||| |||||  ||   ||||| ||||  |: |:|||||:|| || |  |||||
Db         50 NATDDDDLYLEPVQRTAVVLSFRSDKEGTGEKEGTED-SNWAVYFDEKRVLKAGAITLKA 108
```

```
Qy         62 GDNLKIKQ----------NGTNFTYSLKKDLTDLTSVGTEKLSFSAHGNKVNITSDTKGL 111
              ||||||||          | ::|||||||||||||||| |||||| |:||||||||||||
Db        109 GDNLKIKQNTNENTNENTNDSSFTYSLKKDLTDLTSVETEKLSFGANGNKVNITSDTKGL 168

Qy        112 NFAKETAGTNGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAG 171
              ||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db        169 NFAKETAGTNGDPTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAG 228

Qy        172 WNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIK 231
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        229 WNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIK 288

Qy        232 EKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETV 291
              |||||||||| | ||||||||||||||||||||||||||||||||||||||||||||||
Db        289 EKDGKLVTGKGKDENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETV 348

Qy        292 TSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSG 351
              |||| |||||| |||||||||||||| ||||||||||||||||||||||||||||||||
Db        349 TSGTKVTFASGNGTTATVSKDDQGNITVKYDVNVGDALNVNQLQNSGWNLDSKAVAGSSG 408

Qy        352 KVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLS 411
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        409 KVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLS 468

Qy        412 VDGD-ALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGI 470
              || : |||||||  ||||||||||||||||||||||||||||||||||||||||||||||
Db        469 VDDEGALNVGSKDANKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGI 528

Qy        471 AQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHF 530
              |||||||| |||||||||||||||||||||||||||||| |||:||||||||||||||||
Db        529 AQAIATAGLAQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDTGNWVIKGTASGNSRGHF 588

Qy        531 GASASVGYQW 540
              ||||||||||
Db        589 GASASVGYQW 598
```

```
RESULT 15
US-09-377-155-15
; Sequence 15, Application US/09377155
; Patent No. 6197312
; GENERAL INFORMATION:
;  APPLICANT:  PEAK, Ian Richard Anselm
;  APPLICANT:  JENNINGS, Michael Paul
;  APPLICANT:  MOXON, E. Richard
;  TITLE OF INVENTION:  NOVEL SURFACE ANTIGEN
;  FILE REFERENCE:  065064/0128
;  CURRENT APPLICATION NUMBER: US/09/377,155
;  CURRENT FILING DATE:  1999-08-19
;  PRIOR APPLICATION NUMBER:  PCT/AU98/01031
;  PRIOR FILING DATE:  1998-12-14
;  PRIOR APPLICATION NUMBER:  GB 9726398.2
;  PRIOR FILING DATE:  1997-12-12
;  NUMBER OF SEQ ID NOS:  33
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 15
;   LENGTH: 599
;   TYPE: PRT
;   ORGANISM: Neisseria meningitidis
US-09-377-155-15

  Query Match             88.2%;  Score 2430.5;  DB 3;  Length 599;
  Best Local Similarity   88.6%;  Pred. No. 3.1e-180;
  Matches  488;  Conservative   14;  Mismatches   34;  Indels   15;  Gaps    4;

Qy          1 NNEEQEEEYLYLHPVQRTVAVLIVNSDKEGAGEKEKVEENSDWAVYFNEKGVLTAREITLK 60
              : :|:||   | || |:  ||    |||| || |    |  |::|::    |    :|||
Db         53 DEDEEEE---LEPVVRSALVLQFMIDKEGNGENEST-GNIGWSIYYDNHNTLHGATVTLK 108

Qy         61 AGDNLKIKQ----------NGTNFTYSLKKDLTDLTSVGTEKLSFSAHGNKVNITSDTKG 110
              |||||||||          | ::|||||||||||||||| |||||| |:||||||||||||
Db        109 AGDNLKIKQNTNKNTNENTNDSSFTYSLKKDLTDLTSVETEKLSFGANGNKVNITSDTKG 168

Qy        111 LNFAKETAGTNGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNA 170
              ||||||||||||||||||||||||||||||||||||||||||||:||||||||||||||
Db        169 LNFAKETAGTNGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDKKRAASVKDVLNA 228

Qy        171 GWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVI 230
              |||||||||||||||||| ||||||||||||||||||||||||:|||||||||||||||
Db        229 GWNIKGVKPGTTASDNVDFVHTYDTVEFLSADTKTTTVNVESKDNGKRTEVKIGAKTSVI 288

Qy        231 KEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFET 290
              |||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||
Db        289 KEKDGKLVTGKGKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFET 348

Qy        291 VTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSS 350
              |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db        349 VTSGTNVTFASGKGTTATVSKDDQGNITVKYDVNVGDALNVNQLQNSGWNLDSKAVAGSS 408

Qy        351 GKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTL 410
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        409 GKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTL 468

Qy        411 SVDG-DALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAG 469
              |||   |||||||  ||||||||||||||||||||||||||||||||||||||||||||
Db        469 SVDDKGALNVGSKDANKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAG 528

Qy        470 IAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGH 529
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        529 IAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGH 588

Qy        530 FGASASVGYQW 540
              |||||||||||
Db        589 FGASASVGYQW 599
```

Search completed: October 6, 2003, 09:35:53
Job time : 16.4407 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: October 6, 2003, 09:13:35 ; Search time 15.4576 Seconds
(without alignments)
3359.577 Million cell updates/sec

Title: US-09-771-382-33
Perfect score: 2757
Sequence: 1 NNEEQEEYLYLHPVQRTVAV.....TASGNSRGHFGASASVGYQW 540

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_76:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2727	98.9	591	2	G81133	adhesin NMB0992 [i
2	2212	80.2	592	2	A81888	probable surface f
3	451	16.4	298	2	I64138	adhesin homolog HI
4	393	14.3	2059	2	D82671	surface protein XF
5	369.5	13.4	1107	2	AC0976	probable autotrans
6	365.5	13.3	1190	2	A82615	surface protein XF
7	358	13.0	1588	2	A86036	probable adhesin Z
8	358	13.0	1588	2	H91188	probable adhesin E
9	331.5	12.0	658	2	AH0110	probable surface p
10	228.5	8.3	1536	2	A43855	high-molecular-wei
11	212	7.7	2020	2	C48399	ABC-type transport
12	209	7.6	2273	2	T09083	hemagglutinin/hemo
13	207	7.5	1910	2	AF0394	probable adhesin h
14	206.5	7.5	4919	2	T31105	hypothetical prote
15	203	7.4	936	2	I40711	sapB protein - Cam
16	200.5	7.3	1477	2	B43855	high-molecular-wei
17	199.5	7.2	5291	2	F90696	hypothetical prote
18	198.5	7.2	1109	2	A56143	surface-array prot
19	195	7.1	585	2	F90961	flagellin [importe
20	195	7.1	585	2	F85809	hypothetical prote
21	195	7.1	1004	2	C82672	surface-exposed ou
22	195	7.1	1461	2	E90696	hypothetical prote
23	194.5	7.1	1035	2	AD3203	autotransporter pr
24	193	7.0	5188	2	B85547	probable RTX famil
25	192.5	7.0	1428	2	AC2224	hypothetical prote
26	192	7.0	1635	2	AI0452	hemolysin [importe
27	191	6.9	1461	2	A85547	hypothetical prote
28	189	6.9	2249	2	A41477	190K surface antig
29	188.5	6.8	4152	2	T31102	filamentous hemagg
30	188	6.8	906	2	C85739	hypothetical prote
31	188	6.8	1011	2	F90879	probable BigA-like
32	187.5	6.8	1651	2	JC1340	outer membrane pro
33	187	6.8	920	2	I40614	surface array prot
34	186.5	6.8	1430	2	AF0351	probable autotrans
35	186.5	6.8	1530	2	AH1396	peptidoglycan anch
36	185	6.7	365	2	AB3486	cell surface prote
37	185	6.7	1091	2	G64964	hypothetical prote
38	185	6.7	2514	2	F81045	hemagglutinin/hemo
39	184	6.7	1361	2	T03415	S-layer protein -
40	183.5	6.7	1608	2	A28182	hemolysin A - Serr
41	183	6.6	2554	2	AB3528	extracellular seri
42	182.5	6.6	1205	2	AH2486	hypothetical prote
43	182.5	6.6	1577	2	A35140	hemolysin A precur
44	182.5	6.6	3013	2	AB0480	probable invasin Y
45	182	6.6	2551	2	B98047	hypothetical prote

ALIGNMENTS

RESULT 1
G81133
adhesin NMB0992 [imported] - Neisseria meningitidis (strain MC58 serogroup B)
C;Species: Neisseria meningitidis
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001
C;Accession: G81133
R;Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B. ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Masignani, V.; Pizza, M. Science 287, 1809-1815, 2000
A;Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.;
A;Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A;Reference number: A81000; MUID:20175755; PMID:10710307
A;Accession: G81133
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-591 <TET>
A;Cross-references: GB:AE002450; GB:AE002098; NID:g7226229; PIDN:AAF41395.1; PID:g722
A;Experimental source: serogroup B, strain MC58
C;Genetics:
A;Gene: NMB0992

Query Match 98.9%; Score 2727; DB 2; Length 591;
Best Local Similarity 99.3%; Pred. No. 4.6e-133;
Matches 536; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

```
Qy          1 NNEEQEEYLYLHPVQRTVAVLIVNSDKEGAGEKEKVEENSDWAVYFNEKGVLTAREITLK 60
              ||||||| ||| |||||||||||||||| |||||||||||||||||||||||||||||||
Db         52 NNEEQEEDLYLDPVQRTVAVLIVNSDKEGTGEKEKVEENSDWAVYFNEKGVLTAREITLK 111

Qy         61 AGDNLKIKQNGTNFTYSLKKDLTDLTSVGTEKLSFSAHGNKVNITSDTKGLNFAKETAGT 120
              ||||||||||||||||||||||||||||||||||||:|||||||||||||||||||||||
Db        112 AGDNLKIKQNGTNFTYSLKKDLTDLTSVGTEKLSFSANGNKVNITSDTKGLNFAKETAGT 171

Qy        121 NGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        172 NGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPG 231

Qy        181 TTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        232 TTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTG 291

Qy        241 KDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        292 KDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFA 351

Qy        301 SGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSP 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        352 SGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSP 411

Qy        361 SKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGDALNVG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        412 SKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGDALNVG 471

Qy        421 SKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLV 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        472 SKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLV 531

Qy        481 QAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        532 QAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 591


RESULT 2
A81888
probable surface fibril protein NMA1200 [imported] - Neisseria meningitidis (strain Z249
C;Species: Neisseria meningitidis
C;Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001
C;Accession: A81888
R;Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel
; Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,
Nature 404, 502-506, 2000
A;Title: Complete DNA sequence of a serogroup A strain of Neisseria menigitidis Z2491.
A;Reference number: A81775; MUID:20222556; PMID:10761919
A;Accession: A81888
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-592 <PAR>
A;Cross-references: GB:AL162755; GB:AL157959; NID:g7379742; PIDN:CAB84461.1; PID:g737989
A;Experimental source: serogroup A, strain Z2491
C;Genetics:
A;Gene: NMA1200

  Query Match             80.2%;  Score 2212;  DB 2;  Length 592;
  Best Local Similarity   82.9%;  Pred. No. 1.4e-106;
  Matches  456;  Conservative   28;  Mismatches   46;  Indels   20;  Gaps    8;

Qy          1 NNEEQEEYLYLHPVQRTVAVLIVNSDKEGAGEKEKVE-ENSDWAVYFNEKGVLTAREITL 59
              : :|:||   |  |||:| |  : :  ||:|| | :    :: :  | :  ::    :||
Db         53 DEDEEEE---LESVQRSV-VGSIQASMEGSGELETISLSMTNDSKEFVDPYIV----VTL 104

Qy         60 KAGDNLKIKQ------NGTNFTYSLKKDLTDLTSVGTEKLSFSAHGNKVNITSDTKGLNF 113
              ||||||||||      | ::|||||||||| | :| |||||| |:| |||| ||||||||
Db        105 KAGDNLKIKQNTNENTNASSFTYSLKKDLTGLINVETEKLSFGANGKKVNIISDTKGLNF 164

Qy        114 AKETAGTNGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWN 173
              |||||||||||||||||||||||||  : |:     :  |     ||||:|||||||||
Db        165 AKETAGTNGDTTVHLNGIGSTLTDTLAGSSASHVDAGNQST--HYTRAASIKDVLNAGWN 222

Qy        174 IKGVKPGTTA--SDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIK 231
              ||||| |:|   |:|||||||||||||||||||||||||||||||||:||||||||||||
Db        223 IKGVKTGSTTGQSENVDFVRTYDTVEFLSADTKTTTVNVESKDNGKRTEVKIGAKTSVIK 282

Qy        232 EKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETV 291
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        283 EKDGKLVTGKGKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETV 342

Qy        292 TSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSG 351
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        343 TSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSG 402

Qy        352 KVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLS 411
              |||||||||||||||||||||||:|||||||||||| ||||||||||||||||||||||||
Db        403 KVISGNVSPSKGKMDETVNINAGNNIEISRNGKNIDIATSMAPQFSSVSLGAGADAPTLS 462

Qy        412 VDGD-ALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGI 470
              || : |||||||  ||||||||||||||||||||||||||||||||||||||||||||||||
Db        463 VDDEGALNVGSKDANKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGI 522

Qy        471 AQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHF 530
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        523 AQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHF 582

Qy        531 GASASVGYQW 540
              ||||||||||
Db        583 GASASVGYQW 592


RESULT 3
I64138
adhesin homolog HI1732 - Haemophilus influenzae (strain Rd KW20)
C;Species: Haemophilus influenzae
C;Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 24-Oct-1997
C;Accession: I64138
R;Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage
; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman
, D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M.
Science 269, 496-512, 1995
A;Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Vente
A;Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
A;Reference number: A64000; MUID:95350630; PMID:7542800
A;Accession: I64138
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-298 <TIGR>
A;Cross-references: GB:U32846; GB:L42023; NID:g1574588; PID:g1574589; TIGR:HI1732

  Query Match             16.4%;  Score 451;  DB 2;  Length 298;
  Best Local Similarity   48.2%;  Pred. No. 2e-16;
  Matches  106;  Conservative   31;  Mismatches   67;  Indels   16;  Gaps    6;

Qy         47 NEKGVLTARE-ITLKAGDNLKIK--QNGTNFTYSLKKDL-------TDLTSVGTEKLSFS 96
              |:   | | : :||||| ||| |  | | : |::| |||       :|  ::|    :
Db         76 NKNQALKAGDTLTLKAGKNLKAKLDQGGKSVTFALAKDLDVKTAKVSDTLTIGGNTPAAG 135

Qy         97 AHGNKVNITSDTKGLNFAKETAGTNGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDD 156
                 ||:|||   ||  ||   |||||| ||||: ||| |   ||||:|:||  : :|
Db        136 GATPKVSITSTADGLKLAK---GTNGDTAVHLNGLASTLPDVTTNTGASTSVT-FSPSDI 191

Qy        157 EKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNG 216
              || |||::|||||||||||| |     ::||| |  || |||:: |  |  | : :|:||
Db        192 EKTRAATIKDVLNAGWNIKGAKVAGGNTENVDLVAGYDNVEFITGDKNTLDVVLTAKENG 251

Qy        217 KKTEVKIGAKTSVIKEKDGKLVTGKD--KGENGSSTDEGE 254
              | ||||   ||||||: :|||:|||       |::|:  |
Db        252 KTTEVKFTPKTSVIKDNNGKLLTGKQLKDANTGTATNATE 291


RESULT 4
D82671
surface protein XF1529 [imported] - Xylella fastidiosa (strain 9a5c)
C;Species: Xylella fastidiosa
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C;Accession: D82671
R;anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Seq
Nature 406, 151-157, 2000
A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A;Reference number: A82515; MUID:20365717; PMID:10910347
A;Note: for a complete list of authors see reference number A59328 below
A;Accession: D82671
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-2059 <SIM>
A;Cross-references: GB:AE003982; GB:AE003849; NID:g9106554; PIDN:AAF84338.1; GSPDB:GN
A;Experimental source: strain 9a5c
R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer
as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Fr
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; La
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins
A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.
, F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri,
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawa
A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silv
M.; Tsuhako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.
A;Reference number: A59328
```

A;Contents: annotation
C;Genetics:
A;Gene: XF1529

Query Match 14.3%; Score 393; DB 2; Length 2059;
Best Local Similarity 23.9%; Pred. No. 2.2e-12;
Matches 170; Conservative 95; Mismatches 217; Indels 230; Gaps 30;

```
Qy         21 LIVNSDKEGAGEKEKVEENSDWAVYFNE-KGVLTARE----ITLKAGDNLKIKQNGT--- 72
Db       1386 VVINNGGKISGVTAGTEETD--AVNFSQLKSISTAVDQGWTLTASGANGSKVASGGTVDL 1443

Qy         73 -----NFTYSLKKDLTDLT----------SVG----------------TEKLSFSA 97
Db       1444 KNTDGNLTISKSGDSNDVVFNLSEDLKEKSITVGNTQLDKDGVKVSSNVLLDSNELVITS 1503

Qy         98 HGNKVNITSDTKGLNFAKETA----GTNGDTTVHLNGI------GSTLTDTLLNTGA--T 145
Db       1504 HSSTSSVKTLANGESVVNRTVVNGDGVNIDDVVVVNDLGLSIVGGASLTLSGINAGSHKI 1563

Qy        146 TNVT-----NDNVTDDEKKRAASVKDVLNAGWNI--------KGVKPGTTASDNVD---- 188
Db       1564 TNVTAGTEDTDAVNFSQLK---SVSEAVDKGWTLTASGANGSKVVSGGTVDLKNTDGNLA 1620

Qy        189 ---FVRTYDTVEFLSADTK-------TTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLV 238
Db       1621 ISKSGDSNDVVFNLSKDFKVDEVTAGNTVVNTDGVKVG--SDVSLGAMGLFIANGPSVTA 1678

Qy        239 TGKDKGEN-------GSSTDEGEGLVTAKEVIDAVNKAGWRMKTTT----------ANGQ 281
Db       1679 SGFNAGDKVISHVAVGMADTDAVNVSQLKQAVQSVTVKATRYYSTNDGGTQGGNYDGDGA 1738

Qy        282 TGQ---ADKFETVTSGTNV------TFASGKGTTA------------------------ 307
Db       1739 TGSKAIAAGVGTQASGEGAAAVGSGAAASGKGSTAIGRNAIASADGSVALGDGAKDGGRG 1798

Qy        308 -----------------TVS-----KDDQGNITVMYDVNVG-DALNVNQL----QNSGWN 340
Db       1799 AESYTGKYSGVQNNTVGTVSVGDAAKGETRSISNVADAKEAMDAVNLRQLDAVAQKSNLQ 1858

Qy        341 LDS----------------------------KAVAGSSGKVISGNVSPSKGKMDETVNIN 372
Db       1859 TDDMRHEINNIEDVFKITKGDSASSVKGMGVNAMAIGTNAAVSGTESVALGK---NTNVS 1915

Qy        373 AGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGDALNVGSKKDNKPVRITN 432
Db       1916 ADNAVAI-GNG-------SVADRANSVSVGSG--------------GSER-----QVTN 1947

Qy        433 VAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNAR----AGIAQAIATAGLVQAYLPGKS 488
Db       1948 VAAGTADTDAVNVSQLNQGLITAKQYTDGMVGNLRRETSGGVAAAIATANLPQAYVQGRG 2007

Qy        489 MMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 540
Db       2008 MTSVGVSSYQGQSAIAVGVSAVSESGHWVFKFSGSANTRSHVGVGAGVGYQW 2059
```

RESULT 5
AC0976
probable autotransporter sapB [imported] - Salmonella enterica subsp. enterica serovar T
C;Species: Salmonella enterica subsp. enterica serovar Typhi
A;Note: this species has also been called Salmonella typhi
C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C;Accession: AC0976
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A;Reference number: AB0502; MUID:21534947; PMID:11677608
A;Accession: AC0976
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1107 <PAR>
A;Cross-references: GB:AL513382; PIDN:CAD03303.1; PID:g16504923; GSPDB:GN00176
C;Genetics:
A;Gene: sapB

Query Match 13.4%; Score 369.5; DB 2; Length 1107;
Best Local Similarity 21.6%; Pred. No. 1.6e-11;
Matches 164; Conservative 100; Mismatches 211; Indels 283; Gaps 28;

```
Qy         34 EKVEENSDWAVYFNEKGVLTAREITLKAGDNLKIKQNGT---NFTYSLKKDLTDLTSVGT 90
Db       382 EKVDQNT----------------TDIAANTTNITQNSTAIENLNTSVSDINTSITGLTD 424

Qy         91 EKL-------SFSAH--GNKVNITSDTKGLNFAKETAGTNG----DTTVHLNGIGSTLTD 137
Db       425 NALLWDEDTGAFSANHGGSTSKITNVAAGALSEDSTDAVNGSQLYETNQKVDQNTSAIAD 484

Qy        138 TLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFV---RTYD 194
Db       485 --INT-SITNLGTDALSWDDEEGAFSASHGTSGTNKITNVAAGEIASDSTDAINGSQLYE 541

Qy        195 T----------VEFLSADTKTTTVN--------VESKDNGKKTEVKIGAKTSVIKEKDGK 236
Db       542 TNMLISQYNESISQLAGDTSETYITENGTGVKYIRTNDNGLEGQ-------DAYATGNGA 594

Qy        237 LVTGKDK----------GENGSSTDEG-------------------------EGLV---- 257
Db       595 TAVGYDAVASGAGCLALGQNSSSSIEGSIALGSGSTSNRAITTGIRETSATSDGVVIGYN 654

Qy        258 -TAKEVIDAVNKA----GWRMKTTTANGQTGQ-----------------ADKFETVTSG 294
Db       655 TTDRELLGALSLGTDGESYRQITNVADGSEAQDAVTVRQLQNAIGAVTTTPTKYYHANST 714

Qy        295 TNVTFASGKGTTA----TVSKDDQG-----NITVMYDVNVGDALNVN------------ 332
Db       715 EEDSLAVGTDSLAMGAKTIVNADAGIGIGLNTLVMADAINGIAIGSNARANHANSIAMGN 774

Qy        333 --------QLQNSGWNLDSK-------AVAGSSGKVISGNVSPSKGKMDETVNINAG--- 374
Db       775 GSQTTRGAQTDYTAYNMDTPQNSVGEFSVGSEDGQRQITNVAAGSADTDA---VNVGQLK 831

Qy        375 -NNIEITRNGKNI----------------------DIATSMTPQF-------------- 396
Db       832 VTDAQVSRNTQSITNLNTQVSNLDTRVTNIENGIGDIVTTGSTKYFKTNTDGADANAQGA 891

Qy        397 SSVSLGAGADAP----------TLSVDGDALNVGSKKDNKPVRITNVAPGVKEGDVTNVA 446
Db       892 DSVAIGSGSIAAAENSVALGTNSVADEANTVSVGSSTQQR--RITNVAAGVNNTDAVNVA 949

Qy        447 QLK-----------------------------------------------GVAQ------ 453
Db       950 QLKASEAGSVRYETNADGSVNYSVLNLGDGSGGTTRIGNVSAAVNDTDAVNYAQLKRSVE 1009

Qy        454 -----------NLNNRIDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAG 502
Db      1010 EANTYTDQKMGEMNSKIKGIENKMSGGIASAMAMAGLPQAYAPGANMTSIAGGTFNGESA 1069

Qy        503 YAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 540
Db      1070 VAIGVSMVSESGGWVYKLQGTSNSQGDYSAAIGAGFQW 1107
```

RESULT 6
A82615
surface protein XF1981 [imported] - Xylella fastidiosa (strain 9a5c)
C;Species: Xylella fastidiosa
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C;Accession: A82615
R;anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Seq
Nature 406, 151-157, 2000
A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A;Reference number: A82515; MUID:20365717; PMID:10910347

```
A;Note: for a complete list of authors see reference number A59328 below
A;Accession: A82615
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1190 <SIM>
A;Cross-references: GB:AE004017; GB:AE003849; NID:g9107083; PIDN:AAF84783.1; GSPDB:GN001
A;Experimental source: strain 9a5c
R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H
as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laign
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E
A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
, F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
M.; Tsuhako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A;Reference number: A59328
A;Contents: annotation
C;Genetics:
A;Gene: XF1981

  Query Match             13.3%;  Score 365.5;  DB 2;  Length 1190;
  Best Local Similarity   27.0%;  Pred. No. 2.9e-11;
  Matches  150;  Conservative   68;  Mismatches  209;  Indels  129;  Gaps   24;

Qy         43 AVYFNEKGVLTAREITLKAGDNLKIKQNGTNFTYSLKKDLTDLTSV-GTEKLSFSAHGNK 101
Db        706 AVNFSQ---LQAVSSTASKGWNL--LASGANSSNVAPGESVDLKNTDGNIVISKESGSND 760

Qy        102 V--NITSDTKGLNFAKETAG-----TNGDTT---VHLNGIGSTLTDTLLNTGATTNVTND 151
Db        761 VLFNLSS---SLKLDKLTVGDTVMTTNGVTVGSGVTLGSMGLVITDGPSVTSSGINAGSQ 817

Qy        152 NVTDDEKKRAASVKDV--------LNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTK 204
Db        818 KITN----VAAGTADTDAVNLSQLNTAMAGSGAK-------SVHYYSTYD------GGTQ 860

Qy        205 TTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKDKGENG-SSTDEGEGLVTAKEVI 263
Db        861 GGNYNGDGATGTRSIAVGVGTLASA----EGATAVGSGAAASGKGSTAIGRNAVASADGS 916

Qy        264 DAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDV 323
Db        917 VALGD-GAKDGARGAESYTGKYSGLQNNTVGTVSVGDASKGETRTVS-----NVADAKEA 970

Qy        324 NVGDALNVNQL----QNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEI 379
Db        971 T--DAVNLRQLDRVAQDANRYVDNKIESLSEGQTF--------------VKVNSLNN--- 1011

Qy        380 TRNGKNIDIATSMTPQFSSVSLGAGADAPTLSV-----DGDALNVGSK----KDNKPV-- 428
Db       1012 -----------SATP------IAAGVDATAIGVGATASGADSIAMGNKASASADNAVAIG 1054

Qy        429 --------------------RITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNAR- 467
Db       1055 NHSVADRANTVSVGSAGSERQVTNVAAGTADTDAVNVSQLNQGLITAKQYTDGVVGSLRR 1114

Qy        468 ---AGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASG 524
Db       1115 DTDGGVAAAIATANLPQAYIPGRGMTSVGVSSYRGQSAIAVGVSSVSESGRWVFKFSGSA 1174

Qy        525 NSRGHFGASASVGYQW 540
Db       1175 NTRSQVGIGAGVGYQW 1190


RESULT 7
A86036
probable adhesin Z5029 [imported] - Escherichia coli (strain O157:H7, substrain EDL933)
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C;Accession: A86036
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apoda
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: A86036
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1588 <STO>
A;Cross-references: GB:AE005174; NID:g12518349; PIDN:AAG58749.1; GSPDB:GN00145; UWGP:
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: Z5029

  Query Match             13.0%;  Score 358;  DB 2;  Length 1588;
  Best Local Similarity   24.5%;  Pred. No. 1e-10;
  Matches  144;  Conservative   79;  Mismatches  228;  Indels  136;  Gaps   22;

Qy         70 NGTNFTYSLKKDLTDLTSVGTEKLSFSAHGNKVN-ITSDTKGLNFAKETAGTNGDTTVHL 128
Db       1022 NATNMMIEQNTQIINQLAGNTDATYIQENGAGINYVRTNDDGLAFNDASAQGVGATAIGY 1081

Qy        129 NGI----------------------GSTLTDTLLNTGA-TTNVTNDNV------TDDE 157
Db       1082 NSVAKGDSSVAIGQGSYSDVDTGIALGSSSVSSRVIAKGSRDTSITENGVVIGYDTTDGE 1141

Qy        158 KKRAASVKD------VLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVE 211
Db       1142 LLGALSIGDDGKYRQIIN-------VADGSEAHDAVT-VRQLQNAIGAVATTPTKYFHAN 1193

Qy        212 SKDNGK----KTEVKIGAKTSVIKEK----------DGKLVTGKDKGE------------ 245
Db       1194 STEEDSLAVGTDSLAMGAKTIVNGDKGIGIGYGAYVDANALNGIAIGSNAQVIHVNSIAI 1253

Qy        246 -NGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQ-------TGQADKFETVTSGTNV 297
Db       1254 GNGSTTTRGAQTNYTAYNMDAPQNSVGEFSVGSADGQRQITNVAAGSAD-----TDAVNV 1308

Qy        298 TFASGKGTTATVSKDDQGNITVM----------YDVNVGDAL----------NVNQL 334
Db       1309 --GQLKVTDAQVSQNTQ-SITNLDNRVTNLDSRVTNIENGIGDIVTTGSTKYFKTNTDGV 1365

Qy        335 QNSGWNLDSKAVAGSSGKVISGNVSPSKGKM---DETVNINAGNNIEITRN---GKN--- 385
Db       1366 DASAQGKDSVAIGSGSIAAADNSVALGTGSVATEENTISVGSSTNQRRITNVAAGKNATD 1425

Qy        386 -IDIATSMTPQFSSVSLGAGADAPTLSVDGDALNVGSKKDNKPVRITNVAPGVKEGDVTN 444
Db       1426 AVNVAQLKSSEAGGVRYDTKADG---SIDYSNITLGG-GNGGTTRISNVSAGVNNNDVVN 1481

Qy        445 VAQLKGVAQ-----------NLNNRIDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIG 493
Db       1482 YAQLKQSVQETKQYTDQRMVEMDNKLSKTESKLSGGIASAMAMTGLPQAYTPGASMASIG 1541

Qy        494 GGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 540
Db       1542 GGTYNGESAVALGVSMVSANGRWVYKLQGSTNSQGEYSAALGAGIQW 1588


RESULT 8
H91188
probable adhesin ECs4480 [similarity] - Escherichia coli (strain O157:H7, substrain R
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 02-Nov-2001
C;Accession: H91188
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and g
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: H91188
```

```
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1588 <HAY>
A;Cross-references: GB:BA000007; PIDN:BAB37903.1; PID:g13363955; GSPDB:GN00154
A;Experimental source: strain O157:H7, substrain RIMD 0509952
C;Genetics:
A;Gene: ECs4480

  Query Match            13.0%;  Score 358;  DB 2;  Length 1588;
  Best Local Similarity  24.5%;  Pred. No. 1e-10;
  Matches  144;  Conservative   79;  Mismatches  228;  Indels  136;  Gaps   22;

Qy         70 NGTNFTYSLKKDLTDLTSVGTEKLSFSAHGNKVN-ITSDTKGLNFAKETAGTNGDTTVHL 128
              | ||         : :  :  |:      :|  :| : ::  || |   :|   | | :
Db       1022 NATNMMIEQNTQIINQLAGNTDATYIQENGAGINYVRTNDDGLAFNDASAQGVGATAIGY 1081

Qy        129 NGI-----------------------GSTLTDTLLNTGA-TTNVTNDNV------TDDE 157
              | :                      |:::  ::  |:  |::| : |      || |
Db       1082 NSVAKGDSSVAIGQGSYSDVDTGIALGSSSVSSRVIAKGSRDTSITENGVVIGYDTTDGE 1141

Qy        158 KKRAASVKD------VLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVE 211
                | |: |      ::|       |  |: | | |  ||          | | |   :
Db       1142 LLGALSIGDDGKYRQIIN-------VADGSEAHDAVT-VRQLQNAIGAVATTPTKYFHAN 1193

Qy        212 SKDNGK----KTEVKIGAKTSVIKEK----------DGKLVTGKDKGE------------ 245
              | :          : :||||  |  :|          |   : |   |
Db       1194 STEEDSLAVGTDSLAMGAKTIVNGDKGIGIGYGAYVDANALNGIAIGSNAQVIHVNSIAI 1253

Qy        246 -NGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQ-------TGQADKFETVTSGTNV 297
               |||:|  |          :||   :       :|:||         | ||      |   ||
Db       1254 GNGSTTTRGAQTNYTAYNMDAPQNSVGEFSVGSADGQRQITNVAAGSAD-----TDAVNV 1308

Qy        298 TFASGKGTTATVSKDDQGNITVM-------------YDVNVGDAL----------NVNQL 334
                  | | ||:: | :|| :                :  :|| :           | : :
Db       1309 --GQLKVTDAQVSQNTQ-SITNLDNRVTNLDSRVTNIENGIGDIVTTGSTKYFKTNTDGV 1365

Qy        335 QNSGWNLDSKAVAGSSGKVISGNVSPSKGKM---DETVNINAGNNIEITRN---GKN--- 385
               |    || |:   |      :|:   | :   : |:::  :  |      |   |||
Db       1366 DASAQGKDSVAIGSGSIAAADNSVALGTGSVATEENTISVGSSTNQRRITNVAAGKNATD 1425

Qy        386 -IDIATSMTPQFSSVSLGAGADAPTLSVDGDALNVGSKKDNKPVRITNVAPGVKEGDVTN 444
               :::|   : :   |      ||    |:|   : :|   :    ||:||: ||   || |
Db       1426 AVNVAQLKSSEAGGVRYDTKADG---SIDYSNITLGG-GNGGTTRISNVSAGVNNNDVVN 1481

Qy        445 VAQLKGVAQ----------NLNNRIDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIG 493
               ||||   |           ::|::   :     ||| |:|  || ||| || || :||
Db       1482 YAQLKQSVQETKQYTDQRMVEMDNKLSKTESKLSGGIASAMAMTGLPQAYTPGASMASIG 1541

Qy        494 GGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 540
              |||| ||:  |:| | :|  | |: |   | ||:| : |:   | ||
Db       1542 GGTYNGESAVALGVSMVSANGRWVYKLQGSTNSQGEYSAALGAGIQW 1588


RESULT 9
AH0110
probable surface protein (partial) YPO0902 [imported] - Yersinia pestis (strain CO92)
C;Species: Yersinia pestis
C;Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 02-Nov-2001
C;Accession: AH0110
R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,
Nature 413, 523-527, 2001
A;Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A;Reference number: AB0001; MUID:21470413; PMID:11586360
A;Accession: AH0110
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-658 <KUR>
A;Cross-references: GB:AL590842; PIDN:CAC89747.1; PID:g15978974; GSPDB:GN00175
C;Genetics:
A;Gene: YPO0902

  Query Match            12.0%;  Score 331.5;  DB 2;  Length 658;
  Best Local Similarity  25.2%;  Pred. No. 7.7e-10;
  Matches  132;  Conservative   74;  Mismatches  204;  Indels  113;  Gaps   22;

Qy         50 GVLTAREITLKAGDNLKIKQNGTNFTYSLKKDLTDLTSVGTEKLSFSAHGNKVNITSDTK 109
              | :|::  ::  |    |   |   :||         |||   :  :| ||:  ||
Db        217 GAVTSQANSIALG-AASINTVGAQSSYSAYALTAPQASVGELGIG-TALGNR-KITGVAA 273

Qy        110 GLNFAKETAGTNGDTTVHLNGIGSTLTDTLLNTGATTNVTN--DNVTDDEKKRAASVKDV 167
              |      :| ::      |  :|  :         | |:|:    ||  |     |:  :
Db        274 G------SASSDAVNVAQLTAVGDQVQQN------TANITSLGGRVTTIE----GSMASI 317

Qy        168 LNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKT------EV 221
              | |  :|     :|  |:|            :: | :  :   |  :|          |
Db        318 ANGG-GVKYFHANSTQPDSV-----------ASGTNSVAIGPASLASGNAALASGAGAV 364

Qy        222 KIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQ 281
              ||         || :  |:  |:||   :   |                     :  :|
Db        365 AIG--DGAAASADGSVAIGQGSGDNGRGVENYIG----------------KYSNASNTS 405

Qy        282 TGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNL 341
              :|      ||: |   |     | | |||    |       :   ||:|: ||
Db        406 SG------TVSVGNTAT-----GETRTVSNVADG-------LQATDAVNLRQLDG----- 442

Qy        342 DSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNI-EITRNGKNIDI--ATSMTPQFSS 398
                      :|  |:  |||  :   |    :|  : : : :  | |     | |:    :|
Db        443 -----IAASIVVVENNVSGLQNGTDGMFQVNNSSGLAKPSATGANSATGGAGSVASGNNS 497

Qy        399 VSLGAGADAP----------TLSVDGDALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQL 448
               : |:|| |            :::   ::::|||  : :  :||||||  :  |  |  ||
Db        498 TAFGSGAKATAANSAALGANSVADRANSVSVGSVGNER--QITNVAPATQGTDAVNFDQL 555

Qy        449 KGVAQN----LNNRIDNVDGNAR-------AGIAQAIATAGLVQAYLPGKSMMAIGGGTY 497
              | ::       | |   :  : |       ||||  |:: | | | |  | ||  ||  :|
Db        556 KSISNQTNAYTNQRYSELKQDLRKQNSVLSAGIASAMSMASLTQPYTSGSSMTTIGAASY 615

Qy        498 RGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 540
              ||::  ::| ||||| | |: |  || |::| ||    |||||
Db        616 RGQSALSLGVSSISDSGRWVSKLQASSNTQGDFGIGVGVGYQW 658


RESULT 10
A43855
high-molecular-weight surface-exposed protein HMW1 - Haemophilus influenzae
C;Species: Haemophilus influenzae
C;Date: 10-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 08-Oct-1999
C;Accession: A43855
R;Barenkamp, S.J.; Leininger, E.
Infect. Immun. 60, 1302-1313, 1992
A;Title: Cloning, expression, and DNA sequence analysis of genes encoding nontypeable
detella pertussis.
A;Reference number: A43855; MUID:92192797; PMID:1548058
A;Accession: A43855
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1536 <BAR>
A;Cross-references: GB:U08876; GB:M84616; NID:g475770; PIDN:AAA20527.1; PID:g475771
A;Note: sequence inconsistent with the nucleotide translation
A;Note: sequence extracted from NCBI backbone (NCBIN:89235, NCBIP:89239)

  Query Match             8.3%;  Score 228.5;  DB 2;  Length 1536;
  Best Local Similarity  22.1%;  Pred. No. 0.00044;
  Matches  126;  Conservative   83;  Mismatches  215;  Indels  145;  Gaps   26;

Qy         51 VLTAREITLKAGDNLKIKQNGTNFT-YSLKKDLTDLTSVGTEKLSFSAHGN--------- 100
              |: ::   :  | :|: | :|:  | :|::|||| | : |         |
Db        766 VINSKYFNVSTGSSLRFKTSGSTKTGFSIEKDLT-LNATGGNITLLQVEGTDGMIGKGIV 824

Qy        101 -KVNITSDTKGLNFAKETAGT--NGDTTVHLNG----IGSTL----------TDTLLNTG 143
               | ||| :   : |    | |   |: |:: |     |||             | ::|:|
Db        825 AKKNITFEGGNITFGSRKAVTEIEGNVTINNNANVTLIGSDFDNHQKPLTIKKDVIINSG 884

Qy        144 ATT---NVTN--DNVTDDEKKRAASVKDVLNAGWNIKGV-----------KPGTTASD- 185
```

```
Db        885 NLTAGGNIVNIAGNLTVESN---ANFKAITNFTFNVGGLFDNKGNSNISIAKGGARFKDI 941

Qy        186 --------NVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKL 237

Db        942 DNSKNLSITTNSSSTYRTIISGNITNKNGDLNITNE--GSDTEMQIGGDVS---QKEGNL 996

Qy        238 VTGKDK-----------GENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQAD 286

Db        997 TISSDKINITKQITIKAGVDGENSDS-----------DATNNANLTIKT----------- 1034

Qy        287 KFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNV-----NQLQNSGWNL 341

Db       1035 KELKLTQDLNI---SGFNKAEITAKD--GSDLTIGNTNSADGTNAKKVTFNQVKDSKISA 1089

Qy        342 DSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSL 401

Db       1090 DGHKVTLHSKVETSGSNNNTEDSSD--------NNAGLTIDAKNVTVNNNIT-SHKAVSI 1140

Qy        402 GAGADAPTLSVDGDALNVGSKKDNKPVRITNVAPGVK--EGDVTNVAQLKGVAQNLNNRI 459

Db       1141 SATSGEITTKT-GTTINATTGNVEITAQTGSILGGIESSSGSVTLTATEGALA------V 1193

Qy        460 DNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIG------GGTYRGEAGYAIGYSSISDG 513

Db       1194 SNISGNTVTVTANSGALTTLAGSTIKGTESVTTSSQSGDIGGTISGGTVEVKATESLTTQ 1253

Qy        514 GNWIIK----------------GTASGNS 526

Db       1254 SNSKIKATTGEANVTSATGTIGGTISGNT 1282
```

RESULT 11
C48399
ABC-type transport protein ydbA.2 - Escherichia coli (strain K-12)
C;Species: Escherichia coli
C;Date: 19-Nov-1993 #sequence_revision 16-Oct-1998 #text_change 01-Mar-2002
C;Accession: C48399; D64891; H64891
R;Moszer, I.; Glaser, P.; Danchin, A.
Biochimie 73, 1361-1374, 1991
A;Title: Multiple IS insertion sequences near the replication terminus in Escherichia co
A;Reference number: A48399; MUID:92190338; PMID:1665988
A;Accession: C48399
A;Molecule type: DNA
A;Residues: 464-2020 <MOS>
A;Cross-references: GB:D85081; NID:g3041754
A;Experimental source: strain K-12
A;Note: sequence inconsistent with the nucleotide translation
A;Note: sequence extracted from NCBI backbone (NCBIN:88089, NCBIP:88090)
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A;Title: The complete genome sequence of Escherichia coli K-12.
A;Reference number: A64720; MUID:97426617; PMID:9278503
A;Accession: D64891
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-839,'LDLPLYFQTSVIT' <BLA1>
A;Cross-references: GB:AE000237; GB:U00096; NID:g1787665; PIDN:AAC74483.1; PID:g1787668;
A;Experimental source: strain K-12, substrain MG1655
A;Accession: H64891
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 'M',915-2020 <BLA2>
A;Cross-references: GB:AE000237; GB:U00096; NID:g1787665; PIDN:AAC74487.1; PID:g1787672;
A;Experimental source: strain K-12, substrain MG1655
C;Genetics:
A;Gene: ydbA_2
A;Start codon: GTG

Query Match 7.7%; Score 212; DB 2; Length 2020;
Best Local Similarity 24.2%; Pred. No. 0.0044;
Matches 160; Conservative 70; Mismatches 216; Indels 216; Gaps 37;

```
Qy         17 TVAVLIVNSDKEGAGEKEKVEENSDWAVYFNEKGVLTAREITLKAGDNLKIKQNGTNFTY 76

Db        195 TVALEGVSAD--GA---------TKWQYNHNGELVIT--------GDNATVNNNG----- 230

Qy         77 SLKKDLTDLTSVGTE-----------KLSFSAHGNKVNITSDT-----KG--------- 110

Db        231 --KTTVDGKDSTGTEINGNNGKVIQDGDLDVSGGGHGIDITGDSATVDNKGTMTVTDPES 288

Qy        111 ----------LNFAKETAGTNGDTTVHLNG-------IGSTLTDTLLNTGATTNVTNDN 152

Db        289 MGIQIDGDKAIVNNEGESTITNGGTGTQINGDDATANNNGKTTVDGKDSTGTEINGNNGK 348

Qy        153 VTDDEKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEF-LSADTKTTTVNVE 211

Db        349 VIQD------GDLDVSGGGHGI-DITGDSATVDNKGTMTVTDPESIGIQVDGDQAVVNNE 401

Qy        212 SKD-----------NGKKTEVKIGAKTSVIKEKDGKLVTGKD-KGENG----------- 247

Db        402 GESAITNGGTGTQINGDDATANNNGKTTV----DGKDSTGTEIAGNNGKVIQDGDLDVSG 457

Qy        248 ----------SSTDEGEGLVTAKE------VIDA----VNKAGWRMKTTTANGQTGQADK 287

Db        458 GGHGIDITGDSATVDNKGTMTVTDPESIGIQIDGDQAIVNNEG---ESTITNGGTG---- 510

Qy        288 FETVTSGTNVTFASGKGTTATVSKDD-----QGNITVMYDVNVGDALNV----NQLQNSG 338

Db        511 --TQINGNDAT-ANNSGKTTVDGKDSTGTKIAGNIGI---VNLDGSLTVTGGAHGVENIG 564

Qy        339 WN--LDSK-----AVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATS 391

Db        565 DNGTVNNKGDIVVSDTGSIGVLINGEGATVSNTGDVNVS-NEATGFSITTNSGKVSLAGS 623

Qy        392 M-TPQFS----------SVSLGA------GADAPTLSVDGDALNV---GSKKDNKPVRIT 431

Db        624 MQVGDFSTGVDLNGNNNSVTLAAKDLKVVGQKATGINVSGDANTVNITGNVLVDKDKTAD 683

Qy        432 NVA-----PGV------KEGDVTNVAQLKGVAQN-LNNRIDNV-DGNARAGIAQAIATAG 478

Db        684 NAAEYFFDPSVGINVYGSDNNVTLDGKLTVVSDSEVTSRQSNLFDGSAE-------KTSG 736

Qy        479 LV-------------QAYLPGKSMMAIGGGTYRGEAGYAI-------GYSSISDGGNWII 518

Db        737 LVVIGDGNTVNMNGGLELIGEKNALADGSQVTSLRTGYSYTSVIVVSGESSVYLNGDTTI 796

Qy        519 KG 520

Db        797 SG 798
```

RESULT 12
T09083
hemagglutinin/hemolysin-related protein NMB1214 [imported] - Neisseria meningitidis (
N;Alternate names: probable secreted protein
C;Species: Neisseria meningitidis
C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 19-Jan-2001
C;Accession: T09083; B81109
R;Simpson, N.J.; Spratt, B.G.
submitted to the EMBL Data Library, October 1997
A;Reference number: Z16558
A;Accession: T09083
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-2273 <SIM>
A;Cross-references: EMBL:AF030941; NID:g2623257; PID:g2623258
A;Experimental source: strain 44/76
R;Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen,
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.
ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Masignani, V.; Pizza, M.
Science 287, 1809-1815, 2000
A;Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.;
A;Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A;Reference number: A81000; MUID:20175755; PMID:10710307

```
A;Accession: B81109
A;Molecule type: DNA
A;Residues: 1-2273 <TET>
A;Cross-references: GB:AE002469; GB:AE002098; NID:g7226446; PIDN:AAF41596.1; PID:g722645
A;Experimental source: serogroup B, strain MC58
C;Genetics:
A;Gene: pspA; NMB1214

  Query Match             7.6%;  Score 209;  DB 2;  Length 2273;
  Best Local Similarity   20.8%;  Pred. No. 0.0073;
  Matches  132;  Conservative   78;  Mismatches  238;  Indels  188;  Gaps   26;

Qy         47 NEKGVLTAREITLKAGDNLKIKQ---NGTNFTYSLKKDLTDLTSVGTEKLSFSAHGNKVN 103
              |  |::  ||  :    |:|  |    | |   :   |:|:  |:| |         | :
Db       1002 NRGGLIAGREALILNAQNIKNLQGDLQGKNIFAAAGSDITNTGSIGAENALLLKASNNIE 1061

Qy        104 ITSDTKGLNFAKETAGTNGDTT-VHLNG--IGSTLTDT----LLNTGATTNVTNDNVTDD 156
                |:|:     : :    |    ::| |   || | |     :|     || : |  |
Db       1062 SRSETRSNQNEQGSVRNIGRVAGIYLTGRQNGSVLLDAGNNIVLTASELTNQSEDGQT-- 1119

Qy        157 EKKRAASVKDVLNAGWNIKGVKPGTTASDNVDF-----------------VRTYDTVEF- 198
                  ||||| :|:    | : : |  |                   :||   :
Db       1120 ----------VLNAGGDIRSDTTGISRNQNTIFDSDNYVIRKEQNEVGSTIRTRGNLSLN 1169

Qy        199 LSADTKTTTVNVESKDNGKK----TEVKIGAKTSVIKEKDGKLVTGKD------------ 242
                | :    | |:    |      ::|: |  :  : :|    ||:
Db       1170 AKGDIRIRAAEVGSEQGRLKLAAGRDIKVEAGKAHTETEDALKYTGRSGGGIKQKMTRHL 1229

Qy        243 KGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTF--- 299
              | :|| :     | :  ||:|           :  |  |     ||    :::  |:
Db       1230 KNQNGQAV---SGTLDGKEIILVSG------RDITVTGSNIIADNHTILSAKNNIVLKAA 1280

Qy        300 -------------------ASGKGTTATVSKDDQGN--ITVMYDVNVGDALNVNQLQNSG 338
                              : | | ||   || | |   || :  :|  :|| | | ::|
Db       1281 ETRSRSAEMNKKEKSGLMGSGGIGFTAGSKKDTQTNRSETVSHTESVVGSLNGNTLISAG 1340

Qy        339 WNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNN-----IEITRNGKNIDIATSM- 392
               :      ||:     |:|  | ||    ::|:|  |       :      | : :| |:
Db       1341 KHYTQ---TGSTISSPQGDVGISSGK----ISIDAAQNRYSQESKQVYEQKGVTVAISVP 1393

Qy        393 -------------------TPQFSSVSLGAGADAPTLSVD----------------GDAL 417
                                : | |:  | |:|    ||                 |  :
Db       1394 VVNTVMGAVDAVKAVQTVGKSKNSRVNAMAAANALNKGVDSGVALYNAARNPKKAAGQGI 1453

Qy        418 NV---GSKKDNKPVRITNVAPGVKEGDVT---------------------------- 443
              :|    | :|:    ||       |:|| :|
Db       1454 SVSVTYGEQKNTSESRIKGTQ--VQEGKITGGGKVSLTASGAGKDSRITITGSDVYGGKG 1511

Qy        444 ------NVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAYLPGKSM-MAIGGG- 495
                   |  |::   |    | :|      ||:| ||   |:   :  | :     | |
Db       1512 TRLKAENAVQIEAARQTHQERSENKSAGFNAGVAIAI-NKGISFGFTAGANYGKGYGNGD 1570

Qy        496 --TYR-GEAGYAIGYSSISDGGNWIIKGTASGNSRG 528
                 ||    |     ::|  ||: :|||   |  :|
Db       1571 ETAYRNSHIGSKDSQTAIESGGDTVIKG---GQLKG 1603


RESULT 13
AF0394
probable adhesin hmwA [imported] - Yersinia pestis (strain CO92)
C;Species: Yersinia pestis
C;Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 02-Nov-2001
C;Accession: AF0394
R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,
Nature 413, 523-527, 2001
A;Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A;Reference number: AB0001; MUID:21470413; PMID:11586360
A;Accession: AF0394
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1910 <KUR>
A;Cross-references: GB:AL590842; PIDN:CAC92482.1; PID:g15981183; GSPDB:GN00175
C;Genetics:
A;Gene: hmwA

  Query Match             7.5%;  Score 207;  DB 2;  Length 1910;
  Best Local Similarity   23.1%;  Pred. No. 0.0074;
  Matches  126;  Conservative   67;  Mismatches  180;  Indels  172;  Gaps   27;

Qy         49 KGVLTAREITLKAGDNLKIKQNGTNFTYSLKKDLTDLTSVGTEKLSFSAHGNKVNITS-- 106
              ||:  : :  | |  |: :    ||  |:                 |:| ::  | ||||:
Db        803 KGIWISGKTNLNASGNVNLHGVTTNSAYA---------GADAIKISGNSSSNNVNITAGG 853

Qy        107 --DTKGLNFAKETAGT----------NGDTTVHLNGI-GS---TLTDTLLNTGATTNVT 149
                    :|  ||   |           |||  ::: |: ||     | |  |     |:|
Db        854 HISLIAVNGGKEIGSTVSVDYANIIAKNGDFNLNITGMKGSPFNNATITANNISMNGNIT 913

Qy        150 -NDNV--TDDEKKRAASVKDVLNAGWNIKGV----KPGTTASDNVDFVRTYDTVEFLSAD 202
               || |  |:        :|  | :    ||:      | ||::|:  |         ::
Db        914 ANDAVLMTNTFLTAKGDIKTDLTS--PTKGLWFRGNGGMTAANNILLVAN------STSS 965

Qy        203 TKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEV 262
              :|  :|  | :           | ::   ||  :: | |   | : :
Db        966 GETVKINASSSN----------KMNITAGKDISIIAGNSKTATGPNIN---------- 1003

Qy        263 IDAVNKAGWRMKTT----TANGQTGQADKFETVTSGTNVT-----FASGKGTTATVSKDD 313
              |: ||     ::|     | || |       |  || ||:        |    |  :  |:
Db       1004 IENVN-----IETNNGNFTTNGIT------STWLSGVNVSANGVDITSNSTGTGGIVLDN 1052

Qy        314 QGNITVMYDVNV--------GDALNVNQLQNSGWNLDSKAVAGSSGK----VISGNVSPS 361
                :| : |:|         |  :  |   ||  | |   |  |:|:    :| |:   |
Db       1053 TNILTTVGDINTIVTNSSGKGIWIKSNSTLNS--NKDITLVGVSAGQNEGVIIQGSSDAS 1110

Qy        362 KGKMDETVNI----NAGN----------NIEITRNGKNIDIATSMTPQFSSVSLGAGAD 406
              :  :    ||      ||          |: :| :|:||||  |           | | |
Db       1111 RNNISAQGNITLIGKMGNGSGQHSLINLGNVSLTSSGRNIDINGS--------SAGTG-- 1160

Qy        407 APTLSVDGDALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGN- 465
                                  |  |||      |:|:  |: |       | : :: ||
Db       1161 -------------------DVYFTNVELNATAGNVSIYAETKTALSTSLNAVLSLGGNN 1200

Qy        466 ---ARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGY---SSISDGGNWIIK 519
                |: |             :| ||:        :    |  ||:   ||:|  || |:|
Db       1201 SIKAQNG------------WLIGKA--------FNTTQGAGIGFRANSSLSVDGNIILK 1239

Qy        520 GTASG 524
              |   |
Db       1240 GETEG 1244


RESULT 14
T31105
hypothetical protein 2 - Haemophilus ducreyi
C;Species: Haemophilus ducreyi
C;Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999
C;Accession: T31105
R;Ward, C.K.; Lumbley, S.R.; Latimer, J.L.; Cope, L.D.; Hansen, E.J.
J. Bacteriol. 180, 6013-6022, 1998
A;Title: Haemophilus ducreyi secretes a filamentous hemagglutinin-like protein.
A;Reference number: Z20984; MUID:99030326; PMID:9811662
A;Accession: T31105
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-4919 <WAR>
A;Cross-references: EMBL:AF057696; NID:g3929021; PID:g3929023; PIDN:AAC79761.1
C;Genetics:
A;Gene: lspA2

  Query Match             7.5%;  Score 206.5;  DB 2;  Length 4919;
  Best Local Similarity   22.1%;  Pred. No. 0.026;
  Matches  147;  Conservative   73;  Mismatches  213;  Indels  231;  Gaps   32;

Qy         48 EKGVLTAREITLKAGDNLKIKQNGTNFTYSLK-KDLTDLTSVGTEKLSFSAHGNKVNITS 106
              :|  : | ::||   ::: :         :|| | : ||    :|    ||:   :| ::
```

```
Db        477 DKASVLAHKLTLNISNDVSLNNQSKLSANNLKIKKVRDLNLNNSE---LSANNLTLNTSN 533

Qy        107 DTKGLNFAKETAGT---NGDTTVHLNGIGSTLTDTLLNTGATTNVT---------------- 149

Db        534 NITLKNKSKFTAGNMTLNVTNNVTLNN-DSELAANNLTLNVTKNVTLNDASKLSANKLDL 592

Qy        150 --NDNVTDDEKKRAAS-------VKDV---------------LNAGWNIKGVKPGTTASDN 186

Db        593 NVTDNVTLNSKSTLSAGELTFKKVKNVTLNNDSELAANNLSLNASHNVTLNNKSKLSAQK 652

Qy        187 VDF----VRTYDTVEFL--SADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKL--- 237

Db        653 ADIKAVNLTLNDTTELTAKNLDINSTTIT----NNGTIAGIFANITTEKLNNKEKALILA 708

Qy        238 ------------------VTGKDK-----GENGSSTDEGEGLVTAKEVIDAVNKAGWRMK 274

Db        709 EQNLNFTVNGSHYENKGDIVSKDKATVTFSKNSDFTSNGSKLVNAQN----------QLK 758

Qy        275 TTTANGQTGQADKFETVTSGTNVTF-ASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQ 333

Db        759 VNVNNFTISQGD---DITLIGNVTLNASGTFT-------NSGNLTTVKTLDVGDIQNFTN 808

Qy        334 LQN----SGWNLDSKAVAGSSGKVIS----------------------------GN--- 357

Db        809 KGNLTVGEDLHIKSKTKITNDGKLISIKNLNISSEADFINNGTLLGIEALKIATKGNFTN 868

Qy        358 ------------VSPSKGKMDETVN---INAGNNIEITRNGK--NIDIATSMT-PQFSS 398

Db        869 KEKAILASNSLLDISVAEGK--KTFNNGTIESGKNLNITNTGAFLNVDNATIRSFGVLNI 926

Qy        399 VSLGAGADAPTLSVDGDALNVGSKKDNKPVRITNVAPG----------VKEGDVTN--- 444

Db        927 TSTGNVSNNGTL-ISNERLNITS-----AANFTNESNGTVMSNGLLNIIAKQGNITNKNL 980

Qy        445 -----VAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLV------QAYLPG------- 486

Db        981 IASRQQLNLTAVADNITN-----DSNISNKIA-VLHSLGNISLNSKDQVYNLGEIYAGNN 1034

Qy        487 ------------KSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASA 534

Db       1035 ISVKAHQLKNDVKLMGDITTKTKEGQASYKLYQAS----------------NGGHFGNDG 1078

Qy        535 SVGY 538

Db       1079 SSGY 1082
```

RESULT 15
I40711
sapB protein - Campylobacter fetus
C;Species: Campylobacter fetus
C;Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 08-Oct-1999
C;Accession: I40711
R;Dworkin, J.; Tummuru, M.K.; Blaser, M.J.
J. Biol. Chem. 270, 15093-15101, 1995
A;Title: Segmental conservation of sapA sequences in type B Campylobacter fetus cells.
A;Reference number: A56999; MUID:95318069; PMID:7797493
A;Accession: I40711
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-936 <RES>
A;Cross-references: EMBL:U25133; NID:g801999; PIDN:AAA79683.1; PID:g802000
C;Genetics:
A;Gene: sapB

Query Match 7.4%; Score 203; DB 2; Length 936;
Best Local Similarity 22.8%; Pred. No. 0.0048;
Matches 130; Conservative 77; Mismatches 244; Indels 120; Gaps 27;

```
Qy         25 SDKEGAGE-----KEKVEENSDWAVYFNEKGVLTAREITLKAGDNLKIKQNGTNFTYSLK 79

Db        281 SDLETSGDFVFNGYEKVG--------FNVLGDIAS--FATDASKSVNVETTGT------- 323

Qy         80 KDLTDLTSVGTEKLSFSAHGNKVNITSDTK---GLNFAKET---AGTNGDTTVHLNGIGS 133

Db        324 --ITAFTAAGTGKVDVVA-GKITALTADSATSVNLTATNDTITLTSANAATSVNLKTSGA 380

Qy        134 TLTDTLLNTGATTNVTND-----NVTDDEKKRAASVKDVLNAGWNIKGVKPGTTASDNVD 188

Db        381 AKSATITSANAAKNITIDATGVAAVTSATAVENLTVKHATNVTLAGNMDKLATVTLDNAA 440

Qy        189 F-----VRTYDTVEFLSADTKTTTVNVESKD-----NGKKTEVKIGA-----KTSVIKEK 233

Db        441 LTAAIDIKSASTLNLINSSVNGHNISTAAKDVTVHLSGSAAKVKLNTTAATDQTVTLKAN 500

Qy        234 --DGKLVTGKDKGENGSSTDEGEG---LVTAKEVIDAVNKAGWRMKTTTAN------GQ 281

Db        501 ATDNSLEFDSGTAKTTSVTASGSGKTLVIKGAEVETLVN-----IDTTAFNGAADVSFGK 555

Qy        282 TGQADKFETVTSGTNVTFASGKGTTATVSK--DDQGNITVMYDVNVGDALNVNQLQNSGW 339

Db        556 DAQSGKF-SVKTGTGDDKIEFVGTTLTEGSVIDGAGNDTIAMKSAALTSANFTMIKNIE- 613

Qy        340 NL---DSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITR--------------- 381

Db        614 NVAISDAVATADLSSSAFKNIIITTKEAADTTLTINKDQVINFTAADRGSVKLITVKLND 673

Qy        382 -NGKN--IDIATSMTPQFSSVSLGAGADAPTLSVDG--DALNVGS-KKDNKPVRITNVAP 435

Db        674 VTGANDVVKIVLDAAAKDASIALGTEATDKALVIDTGIETLNITSLVKATSPENTANTV- 732

Qy        436 GVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGG 495

Db        733 NAKLTDVTSII----------------IDGDAKITLGHA-GTAGTDYSKVSMIDARAL--- 773

Qy        496 TYRGEAGYAIGYSSISDGGNWIIKGTASGNS 526

Db        774 ----KAGLTFDASAITLGANATIKGGSGADS 800
```

Search completed: October 6, 2003, 09:33:28
Job time : 18.4576 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: October 6, 2003, 09:06:20 ; Search time 8.33898 Seconds
(without alignments)
3045.266 Million cell updates/sec

Title: US-09-771-382-33
Perfect score: 2757
Sequence: 1 NNEEQEEYLYLHPVQRTVAV.....TASGNSRGHFGASASVGYQW 540

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID		Description
1	212	7.7	2003	1	YDBA_ECOLI	P33666	escherichia
2	189	6.9	2249	1	OMPA_RICRI	P15921	rickettsia
3	187.5	6.8	1953	1	BIGA_SALTY	P25927	salmonella
4	185	6.7	1654	1	OMPB_RICRI	Q53047	r outer mem
5	183.5	6.7	1608	1	HLYA_SERMA	P15320	serratia ma
6	182.5	6.6	1039	1	AG43_ECOLI	P39180	escherichia
7	182.5	6.6	1577	1	HLYA_PROMI	P16466	proteus mir
8	181.5	6.6	1645	1	OMPB_RICTY	P96989	r outer mem
9	177.5	6.4	1007	1	Y741_CHLMU	Q9pjt6	chlamydia m
10	177	6.4	917	1	HXA3_HAEIN	P45355	haemophilus
11	177	6.4	1655	1	OMPB_RICCN	Q9kka3	r outer mem
12	176.5	6.4	1569	1	YPJA_ECOLI	P52143	escherichia
13	174	6.3	1567	1	ICEN_XANCT	P18127	xanthomonas
14	173.5	6.3	2334	1	WAPA_BACSU	Q07833	bacillus su
15	173	6.3	933	1	SLAP_CAMFE	P35827	campylobact
16	170.5	6.2	1656	1	OMPB_RICJA	O06653	r outer mem
17	170.5	6.2	2021	1	OMPA_RICCN	Q52657	rickettsia
18	169.5	6.1	928	1	HXA2_HAEIN	P45354	haemophilus
19	169	6.1	918	1	YMJB_CAEEL	P34487	caenorhabdi
20	167	6.1	1643	1	OMPB_RICPR	Q53020	r outer mem
21	163.5	5.9	1861	1	APU_THETU	P38536	t amylopull
22	162	5.9	2660	1	YEEJ_ECO57	Q8x8v7	escherichia
23	161.5	5.9	1286	1	AIDA_ECOLI	Q03155	escherichia
24	161	5.8	550	1	FLIC_SHIFL	Q08860	shigella fl
25	161	5.8	1148	1	ICEK_PSESX	O30611	pseudomonas
26	160	5.8	1300	1	120K_RICRI	P14914	rickettsia
27	159.5	5.8	642	1	FLID_CAMJE	Q9phw6	campylobact
28	159.5	5.8	2358	1	YEEJ_ECOLI	P76347	escherichia
29	158.5	5.7	1025	1	SLAP_CAUCR	P35828	caulobacter
30	157.5	5.7	3178	1	YS89_CAEEL	Q09624	caenorhabdi
31	157	5.7	737	1	ALYS_ENTFA	P37710	enterococcu
32	157	5.7	928	1	PM10_CHLPN	Q9rb65	chlamydia p
33	157	5.7	1153	1	PVDB_PLAKN	P50493	plasmodium
34	156	5.7	497	1	FLIC_ECOLI	P04949	escherichia
35	155	5.6	1196	1	ICEV_PSESX	O33479	pseudomonas
36	155	5.6	1200	1	ICEN_PSESY	P06620	pseudomonas
37	154.5	5.6	507	1	FLIC_SALON	Q06974	salmonella
38	153	5.5	1694	1	IGA0_HAEIN	P44969	haemophilus
39	153	5.5	1702	1	IGA2_HAEIN	P45384	haemophilus
40	152.5	5.5	507	1	FLIC_SALBE	Q06968	salmonella
41	152.5	5.5	537	1	TEE6_STRPY	P18481	streptococc
42	152	5.5	954	1	FLEY_CAUCR	P15345	caulobacter
43	152	5.5	1310	1	VAC3_HELPY	Q48253	helicobacte
44	152	5.5	1325	1	YDEK_ECOLI	P32051	escherichia
45	151.5	5.5	817	1	YG4A_YEAST	P46949	saccharomyc

ALIGNMENTS

RESULT 1
YDBA_ECOLI
ID YDBA_ECOLI STANDARD; PRT; 2003 AA.
AC P33666; P76087; P76088; P76856; P76857; P76859;
DT 01-FEB-1994 (Rel. 28, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein ydbA.
GN YDBA OR B1401/B1405.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=97251357; PubMed=9097039;
RA Aiba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K., Itoh T.,
RA Kasai H., Kashimoto K., Kimura S., Kitakawa M., Kitagawa M.,
RA Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,
RA Nakade S., Nakamura Y., Nashimoto H., Nishio Y., Oshima T., Saito N.,
RA Sampei G., Seki Y., Sivasundaram S., Tagami H., Takeda J.,
RA Takemoto K., Takeuchi Y., Wada C., Yamamoto Y., Horiuchi T.;
RT "A 570-kb DNA sequence of the Escherichia coli K-12 genome
RT corresponding to the 28.0-40.1 min region on the linkage map.";
RL DNA Res. 3:363-377(1996).
RN [3]
RP SEQUENCE OF 464-2003 FROM N.A.
RC STRAIN=K12;
RX MEDLINE=92190338; PubMed=1665988;
RA Moszer I., Glaser P., Danchin A.;
RT "Multiple IS insertion sequences near the replication terminus in
RT Escherichia coli K-12.";
RL Biochimie 73:1361-1374(1991).
CC -!- SIMILARITY: TO S.TYPHIMURIUM ORF NEAR CYSG (AC P25928).
CC -!- CAUTION: THIS IS A CONCEPTUAL TRANSLATION; THE GENE CODING FOR
CC THIS PROTEIN IS INTERRUPTED BY A HYBRID IS2D/IS30C ELEMENT
CC BETWEEN AMINO ACIDS 839 AND 840.
CC ---

```
CC   ----------------------------------------------------------------------
DR   EMBL; AE000237; AAC74483.1; ALT_SEQ.
DR   EMBL; AE000237; AAC74487.1; ALT_SEQ.
DR   EMBL; D90778; BAA15009.1; ALT_SEQ.
DR   EMBL; D90778; BAA18880.1; ALT_SEQ.
DR   EMBL; D90779; BAA18881.1; ALT_SEQ.
DR   EMBL; X62680; -; NOT_ANNOTATED_CDS.
DR   EcoGene; EG11307; ydbA.
KW   Hypothetical protein; Complete proteome.
FT   CONFLICT    489    489       I -> V (IN REF. 2).
FT   CONFLICT    495    495       I -> V (IN REF. 2).
SQ   SEQUENCE   2003 AA;  205949 MW;  B83A12C8B53220EE CRC64;

  Query Match             7.7%;  Score 212;  DB 1;  Length 2003;
  Best Local Similarity   24.2%;  Pred. No. 0.003;
  Matches  160;  Conservative   70;  Mismatches  216;  Indels  216;  Gaps   37;

Qy         17 TVAVLIVNSDKEGAGEKEKVEENSDWAVYFNEKGVLTAREITLKAGDNLKIKQNGTNFTY 76
              |||:  |::|  ||         : |    | : |:|        |||  :  ||
Db        195 TVALEGVSAD--GA---------TKWQYNHNGELVIT--------GDNATVNNNG----- 230

Qy         77 SLKKDLTDLTSVGTE-----------KLSFSAHGNKVNITSDT-----KG--------- 110
                |  :   | |||            |  |  |: ::|| |:     ||
Db        231 --KTTVDGKDSTGTEINGNNGKVIQDGDLDVSGGGHGIDITGDSATVDNKGTMTVTDPES 288

Qy        111 -----------LNFAKETAGTNGDTTVHLNG-------IGSTLTDTLLNTGATTNVTNDN 152
                        :|   |:  ||| |   :||        | |  |   :||   |  |
Db        289 MGIQIDGDKAIVNNEGESTITNGGTGTQINGDDATANNNGKTTVDGKDSTGTEINGNNGK 348

Qy        153 VTDDEKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEF-LSADTKTTTVNVE 211
              |  |          ||   |  |  :   :   ||   :   |     :  |     || |
Db        349 VIQD------GDLDVSGGGHGI-DITGDSATVDNKGTMTVTDPESIGIQVDGDQAVVNNE 401

Qy        212 SKD-----------NGKKTEVKIGAKTSVIKEKDGKLVTGKD-KGENG------------ 247
               :             ||          ||:|  .  |||  ||  :  | ||
Db        402 GESAITNGGTGTQINGDDATANNNGKTTV----DGKDSTGTEIAGNNGKVIQDGDLDVSG 457

Qy        248 ----------SSTDEGEGLVTAKE------VIDA----VNKAGWRMKTTTANGQTGQADK 287
                       |:| : :| :|  :        ||      ||  |    ::|  || ||
Db        458 GGHGIDITGDSATVDNKGTMTVTDPESIGIQIDGDQAIVNNEG---ESTITNGGTG---- 510

Qy        288 FETVTSGTNVTFASGKGTTATVSKDD-----QGNITVMYDVNVGDALNV----NQLQNSG 338
              |  :| : | |:  | |     ||        ||| :   ||:  :| |    : ::| |
Db        511 --TQINGNDAT-ANNSGKTTVDGKDSTGTKIAGNIGI---VNLDGSLTVTGGAHGVENIG 564

Qy        339 WN--LDSK-----AVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATS 391
               |  :::|     :  || | :|:|  :       |  |: |       || |    : :| |
Db        565 DNGTVNNKGDIVVSDTGSIGVLINGEGATVSNTGDVNVS-NEATGFSITTNSGKVSLAGS 623

Qy        392 M-TPQFS----------SVSLGA------GADAPTLSVDGDALNV---GSKKDNKPVRIT 431
              |    ||          ||:| |       |  |  ::| |||  |    |:   :|
Db        624 MQVGDFSTGVDLNGNNNSVTLAAKDLKVVGQKATGINVSGDANTVNITGNVLVDKDKTAD 683

Qy        432 NVA-----PGV------KEGDVTNVAQLKGVAQN-LNNRIDNV-DGNARAGIAQAIATAG 478
              | |     | |       : :||   :|  |: : : :|  |: ||:|          |:|
Db        684 NAAEYFFDPSVGINVYGSDNNVTLDGKLTVVSDSEVTSRQSNLFDGSAE-------KTSG 736

Qy        479 LV-------------QAYLPGKSMMAIGGGTYRGEAGYAI-------GYSSISDGGNWII 518
              ||              :  |: :| |         ||:        | ||:   |:  |
Db        737 LVVIGDGNTVNMNGGLELIGEKNALADGSQVTSLRTGYSYTSVIVVSGESSVYLNGDTTI 796

Qy        519 KG 520
               |
Db        797 SG 798


RESULT 2
OMPA_RICRI
ID   OMPA_RICRI     STANDARD;      PRT;  2249 AA.
AC   P15921;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Outer membrane protein A precursor (190 kDa antigen) (Cell surface
DE   antigen) (rOmpA) (rOmp A).
GN   OMPA.
OS   Rickettsia rickettsii.
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC   Rickettsiaceae; Rickettsieae; Rickettsia.
OX   NCBI_TaxID=783;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=R;
RX   MEDLINE=90354033; PubMed=2117568;
RA   Anderson B.E., McDonald G.A., Jones D.C., Regnery R.L.;
RT   "A protective protein antigen of Rickettsia rickettsii has tandemly
RT   repeated, near-identical sequences.";
RL   Infect. Immun. 58:2760-2769(1990).
CC   -!- FUNCTION: ELICITS PROTECTIVE IMMUNITY.
CC   -!- SUBCELLULAR LOCATION: CELL WALL. THIS BACTERIUM IS COVERED BY A
CC       S-LAYER WITH HEXAGONAL SYMMETRY.
CC   -!- PTM: GLYCOSYLATED (PROBABLE).
CC   -!- SIMILARITY: BELONGS TO THE RICKETTSIAE OMPA/OMPB FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; M31227; AAA26380.1; -.
DR   PIR; A41477; A41477.
DR   InterPro; IPR006315; Autotransport.
DR   InterPro; IPR005546; Autotransporter.
DR   Pfam; PF03797; Autotransporter; 1.
DR   TIGRFAMs; TIGR01414; autotrans_barl; 3.
KW   Antigen; Repeat; Signal; Cell wall; S-layer; Glycoprotein.
FT   SIGNAL       1     28       POTENTIAL.
FT   CHAIN       29   2249       OUTER MEMBRANE PROTEIN A.
FT   DOMAIN     212   1180       13 X APPROXIMATE TANDEM REPEATS.
FT   REPEAT     212    286       A (TYPE I).
FT   REPEAT     287    358       B (TYPE II).
FT   REPEAT     359    430       C (TYPE II).
FT   REPEAT     431    505       D (TYPE II).
FT   REPEAT     506    577       E (TYPE II).
FT   REPEAT     578    652       F (TYPE I).
FT   REPEAT     653    724       G (TYPE II).
FT   REPEAT     725    799       H (TYPE I).
FT   REPEAT     800    874       I (TYPE I).
FT   REPEAT     875    949       J (TYPE I).
FT   REPEAT     950   1021       K (TYPE II).
FT   REPEAT    1022   1093       L (TYPE II).
FT   REPEAT    1094   1165       M (TYPE II).
FT   REPEAT    1166   1180       TYPE I (INCOMPLETE).
SQ   SEQUENCE   2249 AA;  224333 MW;  A9D6646C089DF087 CRC64;

  Query Match             6.9%;  Score 189;  DB 1;  Length 2249;
  Best Local Similarity   22.3%;  Pred. No. 0.049;
  Matches  122;  Conservative   64;  Mismatches  208;  Indels  152;  Gaps   25;

Qy         57 ITLKAGDNLKIKQNGTN---FTYSLKKDLTDLTSVGTEKLS--------FSAHGNKVNIT 105
              :|| || :| :  || |      :       : | :|   |          ||  :|: :
Db        134 VTLNAGKSLTL--NGNNAVAANHGFDAPADNYTGLGNIALGGANAALIIQSAAPSKITLA 191

Qy        106 S--DTKGLNFAKETAGTNGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNV---TDDEKKR 160
                 |  |:   |  |  ||        ||:|     :| || |       |   |  :
Db        192 GNIDGGGIITVKTDAAINGT-------IGNTNALATVNVGAGTATLGGAVIKATTTKLTN 244

Qy        161 AASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTE 220
              ||||  : ||   : |    ||  |||  :     :  :: |   |        |   |
Db        245 AASVLTLTNANAVLTGAIDNTTGGDNVGVLNLNGALSQVTGDIGNT--------NSLAT- 295

Qy        221 VKIGAKT-----SVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKT 275
              : :|| |     :|||    ||        : :|  :    |     ||    |  : |
```

```
Db        296 ISVGAGTATLGGAVIKATTTKLT-------DAASAVKFTNPVVVTGAIDNTGNANNGIVT 348

Qy        276 TTAN----GQTGQADKFETVTSGTNVTFASG---KGTTATVSKDDQGNITVMYDVNVGDA 328
              | |    | | :   || | :    | | | :: |:   :|    | | |
Db        349 FTGNSTVTGNVGNTNALATVNVGAGLLQVQGGVVKANTINLT-DNASAVTFTNPVVVTGA 407

Qy        329 LNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAG-------------- 374
              ::     |:| | ::  |  :    ::|::  :      |||: ||
Db        408 ID-----NTG-NANNGIVTFTGNSTVTGDIGNTNAL--ATVNVGAGTATLGGAVIKATTT 459

Qy        375 ------------------NNIEITRNGKNIDIAT------------SMTPQFSSVSLGA 403
                               |: |  | |: :                 |   :::|:|| |
Db        460 KLTNAASVLTLTNANAVLTGAIDNTTGGDNVGVLNLNGALSQVTGNIGNTNSLATISVGA 519

Qy        404 GADAPTLSVDGDALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVD 463
              |    | :: |  :   : |      :|: |  ||    ||   : |         |||
Db        520 G----TATLGGAVIKATTTK------LTDAASAVK---FTNPVVVTGA-------IDNT- 558

Qy        464 GNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDG------GNWI 517
              |||  ||                :   |  |  |: |     ::|| |      |  :
Db        559 GNANNGI-----------------VTFTGNSTVTGDIGNTNSLATISVGAGTATLGGAV 600

Qy        518 IKGTAS 523
              || | :
Db        601 IKATTT 606


RESULT 3
BIGA_SALTY
ID   BIGA_SALTY     STANDARD;      PRT;  1953 AA.
AC   P25927; P25928; Q9XCQ3;
DT   01-MAY-1992 (Rel. 22, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Putative surface-exposed virulence protein bigA precursor.
GN   BIGA OR STM3478.
OS   Salmonella typhimurium.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Salmonella.
OX   NCBI_TaxID=602;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 14028;
RA   Stojiljkovic I., Valentine P., Heffron F.;
RT   "Salmonella typhimurium rhs homolog.";
RL   Submitted (MAR-1999) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX   MEDLINE=21534948; PubMed=11677609;
RA   McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA   Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA   Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA   Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA   Waterston R., Wilson R.K.;
RT   "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT   LT2.";
RL   Nature 413:852-856(2001).
RN   [3]
RP   SEQUENCE OF 1-765 FROM N.A.
RC   STRAIN=LT2;
RX   MEDLINE=91100301; PubMed=1987123;
RA   Wu J.Y., Siegel L.M., Kredich N.M.;
RT   "High-level expression of Escherichia coli NADPH-sulfite reductase:
RT   requirement for a cloned cysG plasmid to overcome limiting siroheme
RT   cofactor.";
RL   J. Bacteriol. 173:325-333(1991).
CC   -!- CAUTION: Ref.3 sequence differs from that shown due to frameshifts
CC       in positions 414 and 732.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF133696; AAD39458.1; -.
DR   EMBL; AE008859; AAL22340.1; -.
DR   EMBL; M64606; AAA27042.1; ALT_FRAME.
DR   EMBL; M64606; AAA27043.1; ALT_FRAME.
DR   StyGene; SG10437; bigA.
KW   Virulence; Repeat; Signal; Complete proteome.
FT   SIGNAL        1     27       POTENTIAL.
FT   CHAIN        28   1953       PUTATIVE SURFACE-EXPOSED VIRULENCE
FT                                PROTEIN BIGA.
FT   DOMAIN      101    252       15 X 11 AA TANDEM REPEATS.
FT   REPEAT      101    103       1 (INCOMPLETE).
FT   REPEAT      104    113       2 (INCOMPLETE).
FT   REPEAT      114    122       3 (INCOMPLETE).
FT   REPEAT      123    133       4.
FT   REPEAT      134    144       5.
FT   REPEAT      145    155       6.
FT   REPEAT      156    166       7.
FT   REPEAT      167    177       8.
FT   REPEAT      178    188       9.
FT   REPEAT      189    199       10.
FT   REPEAT      200    210       11.
FT   REPEAT      211    221       12.
FT   REPEAT      222    232       13.
FT   REPEAT      233    243       14.
FT   REPEAT      244    252       15 (INCOMPLETE).
FT   CONFLICT    207    207       D -> DRGDDDVTPPDD (IN REF. 1).
FT   CONFLICT    514    514       A -> R (IN REF. 3).
FT   CONFLICT   1698   1698       D -> N (IN REF. 1).
FT   CONFLICT   1795   1798       QYLE -> ITLQ (IN REF. 1).
FT   CONFLICT   1836   1837       SA -> T (IN REF. 1).
SQ   SEQUENCE   1953 AA;  200150 MW;  611B3F1C954D91AE CRC64;

  Query Match             6.8%;  Score 187.5;  DB 1;  Length 1953;
  Best Local Similarity   23.1%;  Pred. No. 0.049;
  Matches  126;  Conservative   63;  Mismatches  195;  Indels  161;  Gaps   29;

Qy         62 GDNLKIKQNGTNFTYSLKKDLTDLTSVGTEKLSFSAHGNKVNITSDTKGLNFAKETAGTN 121
              ||| ||  ||:       |::   | ||         ||   : :|
Db        447 GDNATIKNTGTS-------DISGAGSTGT-----VIDGNNARVNND------------- 480

Qy        122 GDTTVHLNGIGSTLT-DTLL--NTGATT---------NVTNDNVTDDEKKRAASVKDVLN 169
              || |:   | |  :| | ::  | |:||          :  ||             |:|
Db        481 GDMTITDGGTGGHITGDNVVIDNAGSTTVSGADATALYIEGDNAL-----------VIN 528

Qy        170 AG-WNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTS 228
                |  | |  ||   :               | ||     : |      | |
Db        529 EGNQTISGGAVGTRIDGD--------------DAHTTNTGDIAVDGAGSAAVII----- 568

Qy        229 VIKEKDGKLVTGKDKGENGSSTDEGEGLVT--AKEVI--DAVNKAGWRMKTTTANGQTGQ 284
                          |:||| |  |: ||| |  :|     |:|      | |
Db        569 --------------NGDNGSLTQAGDLLVTDGAMGIITYGTGNEA-----KNTGNATVRD 609

Qy        285 ADKFETVTSGTNVTF-------ASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNS 337
              ||    | :|   ||         |  || | ||  |   :|:  |:||   ::|   :
Db        610 ADSVGFVVAGEKNTFKNKGDIDVSLNGTGALVS-GDMSQVTLDGDINV---VSVQDSEGV 665

Qy        338 GWNLDSKAVAGSSGKV-ISGNVSPS---------KGKMDETVNINAGNNIEITRNGK-NI 386
                :  :|:| |  | |:|||: |          |    |  :  ||   :| ||  ||
Db        666 FSSATGVSVSGDSNAVDITGNVNISADYGQDDLAAGAPPLTGVVVGGNGNTVTLNGALNI 725

Qy        387 ---DIATSMTPQFSSVSLGAGADAPTLSVDGDALNVGSKKDNKPVRITNVAPGVKEGDVT 443
                 |:: :    | |    |   : :||  :|:   :|  |:  |:         |:|
Db        726 DDNDLSATGGQYLDVVGLSVTGDDNDVEIDG-GINITHSED--PLDGTS-------ADIT 775

Qy        444 NVAQLKGVAQNLNNRIDNVDGNARAG----IAQAIATAGLV----------QAYLP---- 485
              ::        ||     :| |   |    :|:      |:          :|:|
Db        776 GISVSGNSTVTLNGH-STIDTNTVVGGHVVLARVNNGGSLILGDDSVVDVNVSYIPTGYY 834
```

```
Qy        486 --GKSMMAIGGGTYRGEAG-------YAI----GYSSISDGGNWIIKGTASGNSRGHFGA 532
                  :|| | ||     |       |::      | :|: |: :: |:| :|
Db        835 TYNALLMADGEGTSIENKGDITSHGVYSVIRADNGSEVSNSGDILVYATSSNSSEDRAAI 894

Qy        533 SASVG 537
                : : |
Db        895 TRASG 899


RESULT 4
OMPB_RICRI
ID   OMPB_RICRI     STANDARD;      PRT;  1654 AA.
AC   Q53047;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Outer membrane protein B precursor (168 kDa surface-layer protein)
DE   (Surface protein antigen) (Cell surface antigen 5) (Sca5) (rOmpB)
DE   (rOmp B) [Contains: 120 kDa surface-exposed protein (Surface protein
DE   antigen) (120 kDa outer membrane protein ompB); 32 kDa beta peptide].
GN   OMPB.
OS   Rickettsia rickettsii.
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC   Rickettsiaceae; Rickettsieae; Rickettsia.
OX   NCBI_TaxID=783;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=R;
RX   MEDLINE=92167802; PubMed=1724278;
RA   Gilmore R.D. Jr., Cieplak W. Jr., Policastro P.F., Hackstadt T.;
RT   "The 120 kilodalton outer membrane protein (rOmp B) of Rickettsia
RT   rickettsii is encoded by an unusually long open reading frame:
RT   evidence for protein processing from a large precursor.";
RL   Mol. Microbiol. 5:2361-2370(1991).
RN   [2]
RP   SEQUENCE OF 279-1654 FROM N.A.
RC   STRAIN=R;
RX   MEDLINE=90136087; PubMed=2515418;
RA   Gilmore R.D. Jr., Joste N., McDonald G.A.;
RT   "Cloning, expression and sequence analysis of the gene encoding the
RT   120 kD surface-exposed protein of Rickettsia rickettsii.";
RL   Mol. Microbiol. 3:1579-1586(1989).
CC   -!- FUNCTION: THE 120 kDa SURFACE-EXPOSED PROTEIN IS A MAJOR
CC       STRUCTURAL PROTEIN WHICH MAY PLAY A ROLE AS A RICKETTSIAL
CC       VIRULENCE FACTOR AND/OR IMMUNOGEN DURING INFECTION.
CC   -!- SUBCELLULAR LOCATION: CELL WALL. THIS BACTERIUM IS COVERED BY A S-
CC       LAYER WITH HEXAGONAL SYMMETRY.
CC   -!- SIMILARITY: BELONGS TO THE RICKETTSIAE OMPA/OMPB FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X16353; CAA34403.1; -.
DR   PIR; S18227; S18227.
DR   InterPro; IPR006315; Autotransport.
DR   InterPro; IPR005546; Autotransporter.
DR   Pfam; PF03797; Autotransporter; 1.
DR   TIGRFAMs; TIGR01414; autotrans_barl; 2.
KW   Antigen; S-layer; Cell wall.
FT   CHAIN         1   1333       120 kDa SURFACE-EXPOSED PROTEIN.
FT   CHAIN      1334   1654       32 kDa BETA PEPTIDE.
FT   DOMAIN     1181   1188       POLY-THR.
SQ   SEQUENCE   1654 AA;  168184 MW;  D7AB70FB7087F618 CRC64;

  Query Match             6.7%;  Score 185;  DB 1;  Length 1654;
  Best Local Similarity   22.5%;  Pred. No. 0.054;
  Matches  147;  Conservative   77;  Mismatches  255;  Indels  174;  Gaps   30;

Qy         14 VQRTVAVLIVNSDKEGAGEKEKVEENSDWAVYFNEKGVLTAREITLKAGDNLKIKQNGTN 73
              ||      | |:| :| |   :::  :  :          | ::  |  |||  |  || |
Db        156 VQFNNGAAIDNNDLKGVG---RIDFGAPASTLVFNLANPTTQKAPLILGDNAVI-ANGVN 211

Qy         74 FTYSLKKDLTDLT--SVGTEKLSFSAHGNKVNITSDTKGLNFAKETAGTNGDTTVHLNGI 131
              | ::      ::  |  | |    | |  :   :|    |      |   | ||::  |
Db        212 GTLNVTNGFIQVSNKSFATVKAINIADGQGIIFNTDANNANTLNLQA---GGTTINFTGT 268

Qy        132 GSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVR 191
              |    ||:  |    || |:|              : | |:|||    | :     |
Db        269 DGTGRLVLLSKHAA--ATNFNITG------------SLGGNLKGVIEFNTVA-----VD 308

Qy        192 TYDTVEFLSADTKTTTVN--------VESKDNGKKTEV--KIGAKTSVIKEKDGKLVTGK 241
              |    :|:    | |         | | ||||   :  :: ||  ||:  :    ||:
Db        309 GQLTANAGAANAVIGTNNGAGRAAGFVVSVDNGKVATIDGQVYAKDMVIQSAN---ATGQ 365

Qy        242 -------DKGENGS------------------STDEGEGLVTAKEVIDAV--------- 266
                    | | :|:  .                :|| |  |    :| :|:
Db        366 VNFRHIVDVGADGTTAFKTAASKVTITQDSNFGNTDFG-NLAAQIKVPNAITLTGNFTGD 424

Qy        267 -----NKAGWRMKTTTANG--QTGQADKFETVTSGTNVTFASGKGT-------TATVSKD 312
                  | ||  : |  |||  ::  ||    ||:      |||  |          | :
Db        425 ASNPGNTAG--VITFDANGTLESASADANVAVTNNITAIEASGAGVVQLSGTHAAELRLG 482

Qy        313 DQGNITVMYDVNVGDALN--VNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKM----- 365
              : |:|  : |   |  :|  |||    | | :  :       |:|::  : |
Db        483 NAGSIFKLAD---GTVINGKVNQTALVGGALAAGTITLDGSATITGDIGNAGGAAALQRI 539

Qy        366 --------------------DETVNINA-GNNIEITRNGKNI----DIATSMTPQFSSV 399
                                  |::: | |  |::|    ||    |:| : | |   |
Db        540 TLANDAKKTLTLGGANIIGAGGGTIDLQANGGTIKLTSTQNNIVVDFDLAIA-TDQTGVV 598

Qy        400 SLGAGADAPTLSVDG------------DALNVGSKK---DNKPVRITNVAPGVKEGDV-- 442
              :  :| ||:::|                |:|| |    |  | | :  |  :| |
Db        599 DASSLTNAQTLTINGKIGTIGANNKTLGQFNIGSSKTVLSNGNVAINELVIG-NDGAVQF 657

Qy        443 -------TNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAYLPGK-----SMM 490
                    |      | : : | : |      ||      ||  | :     |     :::
Db        658 AHDTYLITRTTNAAGQGKIIFNPVVNNGTTLAAGTNLGSATNPLAEINFGSKGVNVDTVL 717

Qy        491 AIGGGTYRGEAGYA---------IGYSSISDGGNWIIKGTASGNSRGHFGASA 534
              :| |        ||         :|    : ||  |: ||  |     |   |
Db        718 NVGEGVNL----YATNITTTDANVGSFVFNAGGTNIVSGTVGGQQGNKFNTVA 766


RESULT 5
HLYA_SERMA
ID   HLYA_SERMA     STANDARD;      PRT;  1608 AA.
AC   P15320;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   01-NOV-1990 (Rel. 16, Last annotation update)
DE   Hemolysin precursor.
GN   SHLA.
OS   Serratia marcescens.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Serratia.
OX   NCBI_TaxID=615;
RN   [1]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 31-40.
RC   STRAIN=SN8;
RX   MEDLINE=88257037; PubMed=3290200;
RA   Poole K., Schiebel E., Braun V.;
RT   "Molecular characterization of the hemolysin determinant of Serratia
RT   marcescens.";
RL   J. Bacteriol. 170:3177-3188(1988).
CC   -!- FUNCTION: BACTERIAL HEMOLYSINS ARE EXOTOXINS THAT ATTACK BLOOD
CC       CELL MEMBRANES AND CAUSE CELL RUPTURE BY MECHANISMS NOT CLEARLY
CC       DEFINED.
CC   -!- FUNCTION: CELL-BOUND HEMOLYSIN, WHICH RELEASES HEME-IRON FROM
CC       ERYTHROCYTES BY INTERACTION WITH THE ERYTHROCYTE MEMBRANE. SHLA
CC       REQUIRES SHLB FUNCTION.
CC   -!- SUBCELLULAR LOCATION: Outer membrane.
```

```
CC   -!- SIMILARITY: TO P.MIRABILIS HEMOLYSIN (HPMA).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M22618; AAA50323.1; -.
DR   PIR; A28182; A28182.
KW   Hemolysis; Toxin; Outer membrane; Signal.
FT   SIGNAL        1     30
FT   CHAIN        31   1608       HEMOLYSIN.
SQ   SEQUENCE   1608 AA;  165078 MW;  D669B476FE7DAD51 CRC64;

  Query Match             6.7%;  Score 183.5;  DB 1;  Length 1608;
  Best Local Similarity   21.8%;  Pred. No. 0.062;
  Matches  146;  Conservative   73;  Mismatches  241;  Indels  209;  Gaps   29;

Qy         29 GAGEKEKVEE----------------NSDWAVYFNEKGVLTAREITLKAGDNLKIKQNG 71
              |  ||:|  |                ||| :       : :  |:|||| |:  :
Db        446 GLVEKDKSSERGYQRNHTSSLRTGRWSNSDESESLKASELRSEGELTLKAGRNVSTQGAK 505

Qy         72 TNFTYSLKKDLTDLTSVGTEKLS--------------FSAHGNKVNITSDTKGLNFAKET 117
               :    |  |  :   || :| :                    || |  |: : :: | |
Db        506 VHAQRDLTIDADNQIQVGVQKTANAKAVRDDKTSWGGIGGGDNKNN--SNRREISHASEL 563

Qy        118 AGTNGDTTVHLNG-IGSTLTDTLL--NTGATTNVTNDNVTDDEKKRAASVKDVLNA---- 170
                |:|  |: |||  | |:| :       |     |:  :  |     ::  | ::|
Db        564 --TSGG-TLRLNGQQGVTITGSKARGQKGGEVTATHGGLRID--NALSTTVDKIDARTGT 618

Qy        171 GWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKD-NGKKT----EVKIGA 225
               :||      |::|   |       |   | :||  | |: :  |  | :     |: : :
Db        619 AFNI------TSSSHKADNSYQSSTASELKSDTNLTLVSHKDADVIGSQVASGGELSVES 672

Qy        226 KTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVT----AKEVIDAVNKAGWRM-------K 274
              ||  |  |            |   : || :  :|    |||  |   :|| |:       |
Db        673 KTGNINVK---------AAERQQNIDEQKTALTVNGYAKEAGDKQYRAGLRIEHTRDSEK 723

Qy        275 TT---------------------TANGQTGQADKFETVTSGTNVTFASGKGTTA----- 307
              ||                     | :|    ||| :   ||  |:|  |    ||
Db        724 TTRTENSASSLSGGSVKLKAEKDVTFSGSKLVADKGDASVSGNKVSFLAADDKTASNTEQ 783

Qy        308 ------------------------------------TVSKDDQGNITVMYDVNVGDALN 330
                                                 |   | :||:|    :|  | |
Db        784 TKIGGGFYYTGGIDKLGSGVEAGYENNKTQAQSSKAITSGSDVKGNLT----INARDKLT 839

Qy        331 VNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIAT 390
               |:|      :  ||      :   | :  | |  |||           | |:| :
Db        840 QQGAQHSVGGAYQENAAGVDHLAAADTASTTTTKTDVGVNI----------GANVDYSA 888

Qy        391 SMTPQFSSVSLGA----------GADAPTLSVDGDALNVGSKK--DNKPVRITNVAPG 436
              |   :|  |           |  || : :|  |    |:|   :    :::|  |
Db        889 VTRPVERAVGKAAKLDATGVINDIGGIGAPNVGLDIGAQGGSSEKRSSSSQAVVSSVQAG 948

Qy        437 V----KEGDV----TNVAQLKGVAQNLN----------NRIDNVDGNARAGIAQAIATAG 478
                 :|:|    |     || | ||           || |    : |        :||
Db        949 SIDINAKGEVRDQGTQYQASKG-AVNLTADSHRSEAAANRQDEQSRDTR-------GSAG 1000

Qy        479 LVQAYLPGKSMMAIG----GGTYRGEAGYAIGYSSISDGGNW---------IIKGTASGN 525
               |: |    | : :     ||| |  :  :   :   |  |          | :|||
Db       1001 -VRVYTTTGSDLTVDAKGEGGTQRSNSSASQAVTGSIDAANGINVNVKKDAIYQGTALNG 1059

Qy        526 SRGHFGASA 534
               ||    :|
Db       1060 GRGKTAVNA 1068


RESULT 6
AG43_ECOLI
ID   AG43_ECOLI     STANDARD;      PRT;  1039 AA.
AC   P39180; P75614; P76360; P97241; Q46771;
DT   01-FEB-1995 (Rel. 31, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Antigen 43 precursor (AG43) (Fluffing protein).
GN   FLU OR B2000.
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12 / MG1655;
RX   MEDLINE=97426617; PubMed=9278503;
RA   Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA   Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA   Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA   Mau B., Shao Y.;
RT   "The complete genome sequence of Escherichia coli K-12.";
RL   Science 277:1453-1474(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12;
RX   MEDLINE=97251358; PubMed=9097040;
RA   Itoh T., Aiba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K.,
RA   Kasai H., Kimura S., Kitakawa M., Kitagawa M., Makino K., Miki T.,
RA   Mizobuchi K., Mori H., Mori T., Motomura K., Nakade S., Nakamura Y.,
RA   Nashimoto H., Nishio Y., Oshima T., Saito N., Sampei G., Seki Y.,
RA   Sivasundaram S., Tagami H., Takeda J., Takemoto K., Wada C.,
RA   Yamamoto Y., Horiuchi T.;
RT   "A 460-kb DNA sequence of the Escherichia coli K-12 genome
RT   corresponding to the 40.1-50.0 min region on the linkage map.";
RL   DNA Res. 3:379-392(1996).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ML 308-225;
RA   Henderson I.R., Owen P.;
RL   Submitted (JAN-1997) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   PRELIMINARY SEQUENCE OF 53-78.
RC   STRAIN=ML 308-225;
RX   MEDLINE=89291704; PubMed=2661530;
RA   Caffrey P., Owen P.;
RT   "Purification and N-terminal sequence of the alpha subunit of antigen
RT   43, a unique protein complex associated with the outer membrane of
RT   Escherichia coli.";
RL   J. Bacteriol. 171:3634-3640(1989).
RN   [5]
RP   SEQUENCE OF 53-63.
RC   STRAIN=K12 / EMG2;
RX   MEDLINE=97443975; PubMed=9298646;
RA   Link A.J., Robison K., Church G.M.;
RT   "Comparing the predicted and observed properties of proteins encoded
RT   in the genome of Escherichia coli K-12.";
RL   Electrophoresis 18:1259-1313(1997).
RN   [6]
RP   GENE NAME.
RX   MEDLINE=97257509; PubMed=9103983;
RA   Henderson I.R., Meehan M., Owen P.;
RT   "Antigen 43, a phase-variable bipartite outer membrane protein,
RT   determines colony morphology and autoaggregation in Escherichia coli
RT   K-12.";
RL   FEMS Microbiol. Lett. 149:115-120(1997).
CC   -!- FUNCTION: CONTROLS COLONY FORM VARIATION AND AUTOAGGREGATION. MAY
CC       FUNCTION AS AN ADHESIN.
CC   -!- SUBUNIT: HETERODIMER OF THE PERIPHERAL MEMBRANE PROTEIN (ALPHA
CC       CHAIN) ANCHORED TO THE INTEGRAL OUTER MEMBRANE PROTEIN (BETA
CC       CHAIN).
CC   -!- SUBCELLULAR LOCATION: Outer membrane-associated.
CC   -!- SIMILARITY: TO ADHESIN AIDI-I AND TO BORDETELLA PERTACTIN.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
```

```
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ------------------------------------------------------------------------
DR   EMBL; AE000291; AAC75061.1; ALT_INIT.
DR   EMBL; D90838; BAA15825.1; ALT_INIT.
DR   EMBL; D90839; BAA15832.1; ALT_INIT.
DR   EMBL; U24429; AAB47869.1; -.
DR   HSSP; P07505; 1SRD.
DR   EcoGene; EG12686; flu.
DR   InterPro; IPR006315; Autotransport.
DR   InterPro; IPR005546; Autotransporter.
DR   InterPro; IPR004899; Pertactin.
DR   Pfam; PF03797; Autotransporter; 1.
DR   Pfam; PF03212; Pertactin; 1.
DR   TIGRFAMs; TIGR01414; autotrans_barl; 1.
KW   Outer membrane; Signal; Complete proteome.
FT   SIGNAL        1     52
FT   CHAIN        53    551       ANTIGEN 43 ALPHA CHAIN.
FT   CHAIN       552   1039       ANTIGEN 43 BETA CHAIN.
FT   VARIANT       2      2       K -> N (IN STRAIN ML 308-225).
FT   VARIANT      41     42       SL -> FF (IN STRAIN ML 308-225).
FT   VARIANT      46     46       T -> K (IN STRAIN ML 308-225).
FT   VARIANT     157    157       W -> L (IN STRAIN ML 308-225).
FT   VARIANT     188    188       V -> F (IN STRAIN ML 308-225).
FT   VARIANT     303    305       ATN -> STI (IN STRAIN ML 308-225).
FT   VARIANT     320    320       A -> T (IN STRAIN ML 308-225).
FT   VARIANT     372    372       N -> Q (IN STRAIN ML 308-225).
FT   VARIANT     493    493       E -> V (IN STRAIN ML 308-225).
FT   VARIANT     497    497       S -> N (IN STRAIN ML 308-225).
FT   VARIANT     585    585       H -> Y (IN STRAIN ML 308-225).
FT   VARIANT     709    709       E -> K (IN STRAIN ML 308-225).
FT   VARIANT     721    721       M -> T (IN STRAIN ML 308-225).
FT   VARIANT     751    753       GHL -> SHF (IN STRAIN ML 308-225).
FT   VARIANT     803    803       S -> P (IN STRAIN ML 308-225).
FT   VARIANT     815    815       A -> V (IN STRAIN ML 308-225).
FT   VARIANT     824    824       C -> S (IN STRAIN ML 308-225).
FT   VARIANT     829    835       LNLVHTS -> MNLIYNA (IN STRAIN ML 308-
FT                                225).
FT   VARIANT     845    847       QGT -> LGA (IN STRAIN ML 308-225).
FT   VARIANT     855    855       S -> T (IN STRAIN ML 308-225).
FT   VARIANT     888    888       Q -> L (IN STRAIN ML 308-225).
FT   VARIANT    1025   1025       S -> I (IN STRAIN ML 308-225).
FT   CONFLICT     61     63       ETV -> TTT (IN REF. 5).
SQ   SEQUENCE   1039 AA;  106841 MW;  5170D647C8DEEBE0 CRC64;

  Query Match             6.6%;  Score 182.5;  DB 1;  Length 1039;
  Best Local Similarity   23.7%;  Pred. No. 0.041;
  Matches  116;  Conservative   68;  Mismatches  154;  Indels  151;  Gaps   32;

Qy         98 HGNKVNITSDTKGLNFAK----------------ETAGTNGDTTVHLNGI-----GSTL 135
              | |:: :   | |:  :                  :  ||   |||   |:     | ::
Db         71 HDNQI-VFGTTNGMTISTGLEYGPDNEANTGGQWVQDGGTANKTTVTSGGLQRVNPGGSV 129

Qy        136 TDTLLNTGATTNV----TNDNVTDDEK---KRAASVKDVLN-AGWNIKGVKPGTTASDNV 187
              :J|::: |   ::     |  :   |:   : | :   |:|  || :  |||| |:| |
Db        130 SDTVISAGGGQSLQGRAVNTTLNGGEQWMHEGAIATGTVINDKGWQV--VKPGTVATDTV 187

Qy        188 DFVRT----------YDTVEFLSADTKTTTVNVESKDNGKKTEVKIG-AKTSVIKEKDGK 236
              | |           || :|:  |   ||:|     ||::     | | |:|:
Db        188 --VNTGAEGGPDAENGDTGQFVRGDAVRTTIN----KNGRQIVRAEGTANTTVV------ 235

Qy        237 LVTGKDKGENGSSTDE----GEGLV----TAKEVIDAVNKAGWRM-------KTTTANGQ 281
                | |:  :| : |     |   |    || : :  ||  ||::          || | |
Db        236 -YAGGDQTVHGHALDTTLNGGYQYVHNGGTASDTV--VNSDGWQIVKNGGVAGNTTVN-Q 291

Qy        282 TGQADKFETVTSGTNVTFASG----KGTTATVSKDDQG-NITVMYDVNVGDALNVNQLQN 336
               |:  : :   : ||||   |      | |||:    | |    : |  | | ||  |:|
Db        292 KGRL-QVDAGGTATNVTLKQGGALVTSTAATVT----GINRLGAFSVVEGKADNV-VLEN 345

Qy        337 SGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQF 396
```

```
              |  ||          |::|: : :         :: |  ::: |||      |: |
Db        346 GG-RLD----------VLTGHTATN-------TRVDDGGTLDV-RNG-----GTATT--- 378

Qy        397 SSVSLGAGADAPTLSVDGDALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLN 456
               ||:| |     |  |  |   |:: | |       |  :  |    :   || :  ||
Db        379 --VSMGNGG---VLLADSGAAVSGTRSDGK-------AFSIGGGQADALMLEKGSSFTLN 426

Qy        457 NRIDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNW 516
                          ||      |               : || :    |   | :::::|
Db        427 -----------AGDTATDTT--------------VNGGLFTARGGTLAGTTTLNNGAIL 460

Qy        517 IIKGTASGN 525
               : |    |
Db        461 TLSGKTVNN 469


RESULT 7
HLYA_PROMI
ID   HLYA_PROMI     STANDARD;      PRT;  1577 AA.
AC   P16466;
DT   01-AUG-1990 (Rel. 15, Created)
DT   01-AUG-1990 (Rel. 15, Last sequence update)
DT   01-NOV-1990 (Rel. 16, Last annotation update)
DE   Hemolysin precursor.
GN   HPMA.
OS   Proteus mirabilis.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Proteus.
OX   NCBI_TaxID=584;
RN   [1]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 30-43.
RC   STRAIN=Isolate 477-12;
RX   MEDLINE=90170827; PubMed=2407716;
RA   Uphoff T.S., Welch R.A.;
RT   "Nucleotide sequencing of the Proteus mirabilis calcium-independent
RT   hemolysin genes (hpmA and hpmB) reveals sequence similarity with the
RT   Serratia marcescens hemolysin genes (shlA and shlB).";
RL   J. Bacteriol. 172:1206-1216(1990).
CC   -!- FUNCTION: BACTERIAL HEMOLYSINS ARE EXOTOXINS THAT ATTACK BLOOD
CC       CELL MEMBRANES AND CAUSE CELL RUPTURE BY MECHANISMS NOT CLEARLY
CC       DEFINED.
CC   -!- FUNCTION: CELL-BOUND HEMOLYSIN, WHICH RELEASES HEME-IRON FROM
CC       ERYTHROCYTES BY INTERACTION WITH THE ERYTHROCYTE MEMBRANE. HPMA
CC       REQUIRES HPMB FUNCTION.
CC   -!- SUBCELLULAR LOCATION: Outer membrane.
CC   -!- MISCELLANEOUS: THE CONSERVED AMPHIPATHIC DOMAINS IN SHLA AND HPMA
CC       MAY BE RESPONSIBLE FOR PORE FORMATION.
CC   -!- SIMILARITY: TO S.MARCESCENS HEMOLYSIN (SHLA).
CC   ------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ------------------------------------------------------------------------
DR   EMBL; M30186; AAA25657.1; -.
DR   PIR; A35140; A35140.
KW   Hemolysis; Toxin; Outer membrane; Signal.
FT   SIGNAL        1     29
FT   CHAIN        30   1577       HEMOLYSIN.
SQ   SEQUENCE   1577 AA;  165869 MW;  175975E0C924B2D9 CRC64;

  Query Match             6.6%;  Score 182.5;  DB 1;  Length 1577;
  Best Local Similarity   22.7%;  Pred. No. 0.068;
  Matches  139;  Conservative   71;  Mismatches  229;  Indels  173;  Gaps   30;

Qy         32 EKEKVE------ENSDWAVYFNEKGVLTAREITLKAGDNLKIKQNGTNFTYSLKKDLTDL 85
              |||:::      :  : |     || :|     : ||:|| |  |             |:
Db        383 EKEQIQQIGSQIDAKNNATLTATKGDVTLDAAKINAGNNLAINAN-----------KDI 430

Qy         86 TSVG-TEKLSFSAHGNKVNITS----------DTKGLNFAKETAG-------------- 119
```

```
Db        431 HINGLVEKESRSENGNKRNHTSRLESGSWSNSHQTETLKASELTAGKDLGLDAQGSITAQ 490

Qy        120 -----------TNGDTTVHLN----GIGSTLTDTLL---NTGATTNVTNDN--------- 152

Db        491 GAKLHANENVLVNAKDNINLNVQKTNNDKTVTDNHVMWGGIGGGQNKNNNNQQQVSHATQ 550

Qy        153 VTDDEKKRAASVKDVLNAGWNIKG-----VKPGTTASDNVDFVRTYDTVEFLSAD----- 202

Db        551 LTADGQLLLAADNNVNITGSQVKGNQGAFVK--TTQGDVVIDNALSETISKIDERTGTAF 608

Qy        203 --TKTTTVNVESKDNGKKTEVKIGAKTSVIKEKD----GKLVTGKDK------GE-NGSS 249

Db        609 NITKSSHKNETNKQTSTGSELISDAQLTVVSGNDVNVIGSLIKSADKLGIHSLGDINVKS 668

Qy        250 T-------DEGEGLVT---AKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTF 299

Db        669 AQQVTKIDDEKTSLAITGHAKEVEDKQYSAGFHITHTTNKNTSTETEQANSTISGANVDL 728

Qy        300 ASGKGTTATVS--KDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGN 357

Db        729 QANKDVTFAGSDLKTTAGNASITGD-NVAFVSTENKKQTD--NTDTTISGGFS---YTGG 782

Qy        358 VSPSKGKMDETVNINAGNNIEITRN-GKNIDIATSMTPQFSSVSLGAGA---------- 405

Db        783 VDKVGSKADFQYD-KQHTQTEVTKNRGSQTEVAGDLTITANKDLLHEGASHHVEGRYQES 841

Qy        406 ----------DAPTLSVDGDALNVGSKKD--------NKPVR------ITNVAPGVKEGD 441

Db        842 GENIQHLAVNDSETSKT--DSLNVGIDVGVNLDYSGVTKPVKKAIEDGVNTTKPG-NNTD 898

Qy        442 VTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEA 501

Db        899 LTKKVTARDAIANLAN-LSNLE-TPNVGVEVGIK----------------GGGSQQSQT 939

Qy        502 GYAIGYSSISDG 513

Db        940 DSQAVSTSINAG 951


RESULT 8
OMPB_RICTY
ID   OMPB_RICTY     STANDARD;      PRT;  1645 AA.
AC   P96989;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Outer membrane protein B precursor (168 kDa surface-layer protein)
DE   (Surface protein antigen) (Cell surface antigen 5) (Sca5) (rOmpB)
DE   (rOmp B) [Contains: 120 kDa surface-exposed protein (Surface protein
DE   antigen) (120 kDa outer membrane protein ompB); 32 kDa beta peptide].
GN   OMPB OR SLP.
OS   Rickettsia typhi.
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC   Rickettsiaceae; Rickettsieae; Rickettsia.
OX   NCBI_TaxID=785;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Wilmington;
RX   MEDLINE=94040787; PubMed=8224886;
RA   Hahn M.-J., Kim K.-K., Kim I., Chang W.-H.;
RT   "Cloning and sequence analysis of the gene encoding the crystalline
RT   surface layer protein of Rickettsia typhi.";
RL   Gene 133:129-133(1993).
RN   [2]
RP   PARTIAL SEQUENCE.
RC   STRAIN=Wilmington;
RX   MEDLINE=92114896; PubMed=1370573;
RA   Ching W.M., Carl M., Dasch G.A.;
RT   "Mapping of monoclonal antibody binding sites on CNBr fragments of
RT   the S-layer protein antigens of Rickettsia typhi and Rickettsia
RT   prowazekii.";
RL   Mol. Immunol. 29:95-105(1992).
RN   [3]
RP   IDENTIFICATION OF CLEAVAGE SITE.
RX   MEDLINE=92104668; PubMed=1729180;
RA   Hackstadt T., Messer R., Cieplak W., Peacock M.G.;
RT   "Evidence for proteolytic cleavage of the 120-kilodalton outer
RT   membrane protein of rickettsiae: identification of an avirulent
RT   mutant deficient in processing.";
RL   Infect. Immun. 60:159-165(1992).
CC   -!- FUNCTION: THE 120 kDa SURFACE-EXPOSED PROTEIN IS A MAJOR
CC       STRUCTURAL PROTEIN WHICH MAY PLAY A ROLE AS A RICKETTSIAL
CC       VIRULENCE FACTOR AND/OR IMMUNOGEN DURING INFECTION.
CC   -!- FUNCTION: THE 32 kDa BETA PEPTIDE MAY SERVE AS A MEMBRANE ANCHOR.
CC   -!- SUBCELLULAR LOCATION: CELL WALL. THIS BACTERIUM IS COVERED BY A
CC       S-LAYER WITH HEXAGONAL SYMMETRY.
CC   -!- SIMILARITY: BELONGS TO THE RICKETTSIAE OMPA/OMPB FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; L04661; AAB48987.1; -.
DR   PIR; JN0896; JN0896.
DR   InterPro; IPR006315; Autotransport.
DR   InterPro; IPR005546; Autotransporter.
DR   Pfam; PF03797; Autotransporter; 1.
DR   TIGRFAMs; TIGR01414; autotrans_barl; 1.
KW   Antigen; S-layer; Transmembrane; Cell wall.
FT   CHAIN         1   1353       120 kDa SURFACE-EXPOSED PROTEIN.
FT   CHAIN      1354   1645       32 kDa BETA PEPTIDE.
FT   TRANSMEM   1415   1429       MEMBRANE ANCHOR (POTENTIAL).
FT   CONFLICT    657    657       H -> N (IN REF. 2).
FT   CONFLICT    842    842       V -> I (IN REF. 2).
FT   CONFLICT   1071   1071       G -> A (IN REF. 2).
FT   CONFLICT   1306   1306       G -> S (IN REF. 2).
SQ   SEQUENCE   1645 AA;  169698 MW;  0CB5641C7EB185EE CRC64;

  Query Match             6.6%;  Score 181.5;  DB 1;  Length 1645;
  Best Local Similarity   22.5%;  Pred. No. 0.08;
  Matches  138;  Conservative   67;  Mismatches  220;  Indels  189;  Gaps   32;

Qy         53 TAREITLKAGDNLKIKQNGTNFTYSLKK---DLTDLTSVGTEKLSFSAHGNKVNITSDTK 109

Db        192 TTQEAPLTLGDNAKI-VNGANGILNITNGFVKVSDKTFAGIK---------TINI-GDNQ 240

Qy        110 GLNFAKETAGTN------GDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAAS 163

Db        241 GLMFNTTPDAANALNLQGGGNTINFNGRDG--TGKLVLVSKNGNATEFNVTG-------- 290

Qy        164 VKDVLNAGWNIKGV-KPGTTA----------------SDN-----VDFVRTYDTVEFLS 200

Db        291 -----SLGGNLKGVIEFDTTAAAGKLIANGGAANAVIGTDNGAGRAAGFIVSVD-----N 340

Qy        201 ADTKTTTVNVESKD--------NGKKT-----EVKIGAKTSVIKEKDGKLVTGKDKGENG 247

Db        341 GNAATISGQVYAKDIVIQSANAGGQVTFEHLVDVGLGGKTN-FKTADSKVII----TENA 395

Qy        248 S--STDEG----------EGLVTAKEVIDAVNK---AGWRMKTTTANGQ--TGQADKFET 290

Db        396 SFGSTDFGNLAVQIVVPNNKILTGNFIGDAKNNGNTAG--VITFNANGTLVSGNTDPNIV 453

Qy        291 VTS-------GTNVTFASG-------KGTTATVSKDDQGNITVMYDVNVGDALNVNQLQN 336

Db        454 VTNIKAIEVEGAGIVQLSGIHGAELRLGNAGSIFKLADGTV-INGPVNQNPLVNNNALAA 512

Qy        337 SGWNLDSKAVAGSSGKVISGNVSPSKGKMDET----------------VNINAGNNIEIT 380

Db        513 GSIQLDGSAII--TGDI--GNGAVNAALQDITLANDASKILTLSGANIIGANAGGAIHFQ 568

Qy        381 RNGKNIDIATSM------------TPQFSSVSLGAGADAPTLSVDGD------------A 416
```

```
Db        569 ANGGTIQLTSTQNNILVDFDLDVTTDQTGVVDASSLTNNQTLTINGSIGTIGANTKTLGR 628

Qy        417 LNVGSKK----------------DNKPVRIT-NVAPGVKEGDVTNVAQLKGVAQNLNNRI 459
Db        629 FNVGSSKTILNAGDVAINELVMENDGSVHLTHNTYLITKTINAANQGKIIVAADPINTDT 688

Qy        460 DNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYA---------IGYSSI 510
Db        689 ALADGTNLGSAESPLSNIHFATKAANGDSILHIG----KGVNLYANNITTTDANVGSLHF 744

Qy        511 SDGGNWIIKGTASG 524
Db        745 RSGGTSIVSGTVGG 758


RESULT 9
Y741_CHLMU
ID   Y741_CHLMU     STANDARD;      PRT;  1007 AA.
AC   Q9PJT6;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Hypothetical protein TC0741 precursor.
GN   TC0741.
OS   Chlamydia muridarum.
OC   Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX   NCBI_TaxID=83560;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MoPn / Nigg;
RX   MEDLINE=20150255; PubMed=10684935;
RA   Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA   White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,
RA   Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA   Gwinn M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
RA   Eisen J., Fraser C.M.;
RT   "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
RT   pneumoniae AR39.";
RL   Nucleic Acids Res. 28:1397-1406(2000).
CC   -!- SIMILARITY: BELONGS TO THE CHLAMYDIAL CPN0572/CT456/TC0741
CC       FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE002342; AAF39550.1; -.
DR   PIR; H81670; H81670.
DR   TIGR; TC0741; -.
KW   Hypothetical protein; Signal; Complete proteome.
FT   SIGNAL        1     51       POTENTIAL.
FT   CHAIN        52   1007       HYPOTHETICAL PROTEIN TC0741.
SQ   SEQUENCE   1007 AA;  104006 MW;  842800C0871B1518 CRC64;

  Query Match             6.4%;  Score 177.5;  DB 1;  Length 1007;
  Best Local Similarity   21.8%;  Pred. No. 0.071;
  Matches  140;  Conservative   74;  Mismatches  209;  Indels  219;  Gaps   34;

Qy         39 NSDWAVYFNEKGVLTAREITLKAGDNL-------------------KIKQNGTNFTYSL- 78
Db        309 SSDWG---NRVDPLVSSKAGIESGGHLPSSVIINNKFRTCVAYGPWNPKENGPNYTPSAW 365

Qy         79 -KKDLTDLTSV--GTEKLSFSAHGNKVN----ITSDTKGLNFAKETAGTNGDTTVHLNGI 131
Db        366 RRGHRVDFGKIFDGT------APFNKINWGSSPTPGDDGISFSNETIGSEPFATPPSSPS 419

Qy        132 GSTLTDTLLNTGAT-TNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPGT-TASDNVDF 189

Db        420 QTPVINVNVNVGGTNVNIGDTNVS--------------------KG--SGTPTSSQSVDM 457

Qy        190 VRTYDTVEFLSADTKT---TTVNVESKDNGKKTEVKIGAKTSVIKEKDG----KLVTGKD 242
Db        458 --STDTSDLDTSDIDTNNQTNGDINTNDNSNNVDGSLSDVDSRVEDDDGVSDTESTNGND 515

Qy        243 K-----------------------------GENGSSTDEGEGLVTAKE----VIDAVNK 268
Db        516 SGKTTSTEENGDPSGPDILAAVRKHLDTVYPGENGGST---EGPLPANQNLGNVIHDVEQ 572

Qy        269 AGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDA 328
Db        573 NGSAKETIITPGDTGPTDSSSSVDADADVEDTS--DTDSGIGDDD----------GVSDT 620

Qy        329 LNVNQLQNSGWNLDSKAVAGSSGKVI-------------SGNVSPSKGKMDETVNINAGN 375
Db        621 ESTNG-NNSGKTTSTEENGDPSGPDILAAVRKHLDTVYPGENGGSTEGPL--PANQNLGN 677

Qy        376 NI-EITRNGKNIDIATSMTP-----QFSSVSLGAGADAPTLS------------VDGDAL 417
Db        678 VIHDVEQNGAAQE--TIITPGDTESTDTSSSVNANADLEDVSDADSGFGDDDGISDTEST 735

Qy        418 NVGSKKDNKPV-----------------RITNVAPGVKEGDVTNVAQLKGVAQNLNNRID 460
Db        736 NGNDSGKNTPVGDGGTPSGPDILAAVRKHLDTVYPGENGGSTERPLPAN---QNLGDIIH 792

Qy        461 NVDGNA----------RAGIAQAIATAGLVQAYLP--------------GK----SMMAI 492
Db        793 DVEQNGSAKETVVSPYRGGGGNTSSPIGLA-SLLPATPSTPLMTTPRTNGKAAASSLMIK 851

Qy        493 GGGTYRGEAGYAIGYSSISDGGNWIIKGTAS-----GNSRGH 529
Db        852 GGETQA---------KLVKNGGN--IPGETTLAELLPRLRGH 882


RESULT 10
HXA3_HAEIN
ID   HXA3_HAEIN     STANDARD;      PRT;   917 AA.
AC   P45355;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Heme/hemopexin-binding protein precursor (Heme:hemopexin utilization
DE   protein A).
GN   HXUA.
OS   Haemophilus influenzae.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC   Pasteurellaceae; Haemophilus.
OX   NCBI_TaxID=727;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NTHI N182;
RX   MEDLINE=95115556; PubMed=7815944;
RA   Cope L.D., Thomas S.E., Latimer J.L., Slaughter C.A.,
RA   Mueller-Eberhard U., Hansen E.J.;
RT   "The 100 kDa haem:haemopexin-binding protein of Haemophilus
RT   influenzae: structure and localization.";
RL   Mol. Microbiol. 13:863-873(1994).
CC   -!- FUNCTION: BINDS HEME/HEMOPEXIN COMPLEXES.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U08349; AAA74139.1; -.
KW   Transport; Signal; Repeat.
FT   SIGNAL        1     21       BY SIMILARITY.
FT   CHAIN        22    917       HEME/HEMOPEXIN-BINDING PROTEIN.
```

```
FT   DOMAIN       24     38       3 X 5 AA TANDEM REPEATS.
FT   REPEAT       24     28       1-1.
FT   REPEAT       29     33       1-2.
FT   REPEAT       34     38       1-3.
FT   DOMAIN      111    668       6 X 6 AA APPROXIMATE REPEATS.
FT   REPEAT      111    116       2-1.
FT   REPEAT      203    208       2-2.
FT   REPEAT      277    282       2-3.
FT   REPEAT      399    404       2-4.
FT   REPEAT      624    629       2-5.
FT   REPEAT      663    668       2-6.
FT   DOMAIN      159    170       2 X 6 AA APPROXIMATE TANDEM REPEATS.
FT   REPEAT      159    164       3-1.
FT   REPEAT      165    170       3-2.
SQ   SEQUENCE   917 AA;  100148 MW;  7FF39BB8C046539D CRC64;

  Query Match             6.4%;  Score 177;  DB 1;  Length 917;
  Best Local Similarity   21.1%;  Pred. No. 0.067;
  Matches  118;  Conservative   93;  Mismatches  223;  Indels  124;  Gaps   23;

Qy         14 VQRTVAVLIVNSDKEGAGEKEKVE--ENSDWAVYFNEKGVLTAREITLKAGDNLKI---K 68
              : :|   : :|      |: ::||  : |: :| :|      | :|  |   | |:
Db         59 INQTSDKVQINWHSFDIGQNKEVEFKQPSENSVAYNRVTGGNASQIQGKLTANGKVYLAN 118

Qy         69 QNGTNFTYSLKKDLTDLTSVGTEKLSFSAHGNKVNITSDTKGLNFAKETAGTN-----GD 123
               ||   |  : ::  | :   :    | : |    |  ||     ||    |
Db        119 PNGVIITKGAEINVAGLLATTKDLEQISENSNSYQFTRKTKDRQVVKEGQVINEGEIKAK 178

Qy        124 TTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPGTTA 183
                | |||      | ::| |      | ||   ||    : |  |::|:|        |
Db        179 DFVVLNG------DEVINKG------NINV---EKNSTINGKVYLSSGYNFT-----FTL 218

Qy        184 SDNVDFVRTYDTVEFLSADTKTTTVN--VESKDNGKKTEVKIGAK--------------- 226
              ||:            :|   :  ||   |::: : |  |: : ||
Db        219 SDS-----------GISVALEDNTVRGIVQNEGSIKAGEITLSAKGRKEALDSLVMNNGV 267

Qy        227 --TSVIKEKDGKLVTGKDKGE-NGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTG 283
               : :  ::||:|   |  : |  |  :|| :    ||    :|   ::| |:  |
Db        268 LEATKVSNRNGKVVLSADNVQLNNESNIKGEIVNFGTEVTSNEDK---KLKITSKTGSKV 324

Qy        284 QADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQN-SGWNLD 342
               : |        |:    |:  : |   :::  :    :::|  | |:    :      |
Db        325 TSPKINFKGKSVNINGNFGRENSGTHYNEERKTLNTEVNIDVPGAENIRIADDKDNTETD 384

Qy        343 S-------KAVAGSSGKV---------ISGNVS-------------PSKGKMD-ETVNINA 373
              |        ::  ::|||         ||||::               ::| :: :  :||:
Db        385 SFIQTEALSSLLANNGKVNLKGNDVNISGNINIDSFRGTDSLLKLTNQGHINIDHADINS 444

Qy        374 -GNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGDALNVGSKKDNKPVRIT- 431
              |    :|    ::|  :::|   | ::|| ||     ||: |      |:  |  |
Db        445 KGRLFVVTSLQNDVDFKSNITITDSKINLGNGAMGLGRSVNEDDYATRWKRAEKSQRKKF 504

Qy        432 NVAPGVKEGDVTNVAQLKGVAQNLNNRIDNV---DGNARAGIAQAIATAGLVQAYLPGKS 488
              ||       |: ||         : |::| |    |  :  : : :||         |::
Db        505 NV-------DMRNV---------VFNQVDEVILAGGFEKVNLDKIVAT---------GQT 539

Qy        489 MMAIGGGTYRGEAGYAIG 506
                | ||  |  : |  |
Db        540 NFYIDGGVSRNNSRYEYG 557


RESULT 11
OMPB_RICCN
ID   OMPB_RICCN     STANDARD;      PRT;  1655 AA.
AC   Q9KKA3; Q9KK98; Q9XC45;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Outer membrane protein B precursor (168 kDa surface-layer protein)
DE   (Surface protein antigen) (Cell surface antigen 5) (Sca5) (rOmpB)
DE   (rOmp B) [Contains: 120 kDa surface-exposed protein (Surface protein
DE   antigen) (120 kDa outer membrane protein ompB); 32 kDa beta peptide].
GN   OMPB OR RC1085.
OS   Rickettsia conorii.
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC   Rickettsiaceae; Rickettsieae; Rickettsia.
OX   NCBI_TaxID=781;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Malish 7;
RX   MEDLINE=21442074; PubMed=11557893;
RA   Ogata H., Audic S., Renesto-Audiffren P., Fournier P.-E., Barbe V.,
RA   Samson D., Roux V., Cossart P., Weissenbach J., Claverie J.-M.,
RA   Raoult D.;
RT   "Mechanisms of evolution in Rickettsia conorii and R. prowazekii.";
RL   Science 293:2093-2098(2001).
RN   [2]
RP   SEQUENCE OF 33-1649 FROM N.A.
RC   STRAIN=Indian tick typhus, and Malish 7;
RX   MEDLINE=20393643; PubMed=10939649;
RA   Roux V., Raoult D.;
RT   "Phylogenetic analysis of members of the genus Rickettsia using the
RT   gene coding the outer-membrane protein rOmpB (ompB).";
RL   Int. J. Syst. Evol. Microbiol. 50:1449-1455(2000).
RN   [3]
RP   SEQUENCE OF 353-1655 FROM N.A.
RC   STRAIN=Malish 7;
RA   Stenos J., Walker D.;
RT   "The rickettsial outer membrane protein A and B genes of Rickettsia
RT   australis, the most divergent rickettsia of the spotted fever group.";
RL   Submitted (MAY-1999) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: THE 120 KDA SURFACE-EXPOSED PROTEIN IS A MAJOR
CC       STRUCTURAL PROTEIN WHICH MAY PLAY A ROLE AS A RICKETTSIAL
CC       VIRULENCE FACTOR AND/OR IMMUNOGEN DURING INFECTION (BY
CC       SIMILARITY).
CC   -!- FUNCTION: THE 32 KDA BETA PEPTIDE MAY SERVE AS A MEMBRANE ANCHOR
CC       (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: CELL WALL. THIS BACTERIUM IS COVERED BY A S-
CC       LAYER WITH HEXAGONAL SYMMETRY (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE RICKETTSIAE OMPA/OMPB FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AE008659; AAL03623:1; -.
DR   EMBL; AF123721; AAF34124.1; -.
DR   EMBL; AF123726; AAF34129.1; -.
DR   EMBL; AF149110; AAD39533.1; -.
DR   PIR; E97835; E97835.
DR   InterPro; IPR006315; Autotransport.
DR   InterPro; IPR005546; Autotransporter.
DR   Pfam; PF03797; Autotransporter; 1.
DR   TIGRFAMs; TIGR01414; autotrans_barl; 2.
KW   Antigen; S-layer; Cell wall; Complete proteome.
FT   CHAIN         1   1334       120 kDa SURFACE-EXPOSED PROTEIN.
FT   CHAIN      1335   1655       32 kDa BETA PEPTIDE.
FT   VARIANT      61     61       P -> A (IN STRAIN INDIAN TICK TYPHUS).
FT   VARIANT      75     75       G -> S (IN STRAIN INDIAN TICK TYPHUS).
FT   VARIANT      78     78       K -> N (IN STRAIN INDIAN TICK TYPHUS).
FT   VARIANT     251    251       V -> A (IN STRAIN INDIAN TICK TYPHUS).
FT   VARIANT     413    413       N -> D (IN STRAIN INDIAN TICK TYPHUS).
FT   VARIANT     959    959       I -> V (IN STRAIN INDIAN TICK TYPHUS).
FT   VARIANT     988    988       A -> T (IN STRAIN INDIAN TICK TYPHUS).
FT   VARIANT    1139   1139       R -> L (IN STRAIN INDIAN TICK TYPHUS).
FT   CONFLICT    353    354       KD -> GH (IN REF. 3).
FT   CONFLICT    776    776       F -> S (IN REF. 3).
FT   CONFLICT   1159   1159       E -> D (IN REF. 3).
FT   CONFLICT   1177   1177       G -> S (IN REF. 3).
FT   CONFLICT   1492   1492       H -> R (IN REF. 3).
SQ   SEQUENCE   1655 AA;  168342 MW;  E49E19377D5FCE37 CRC64;
```

```
  Query Match             6.4%;  Score 177;  DB 1;  Length 1655;
  Best Local Similarity   23.9%;  Pred. No. 0.13;
  Matches  142;  Conservative   71;  Mismatches  265;  Indels  116;  Gaps   30;

Qy         14 VQRTVAVLIVNSDKEGAGEKEKVEENSDWAVYFNEKGVLTAREITLKAGDNLK--IKQNG 71
              :|   | :  |    :| |    : :|:  |  ||  | |          | |||  |: | .
Db        256 LQANGATITFNG-TDGTGRLVLLSKNAA-ATDFNVTGSL---------GGNLKGIIEFNT 304

Qy         72 TNFTYSLKKDL-TDLTSVGTEKLSFSAHGNKVNI----TSDTKGLNFAK----ETAGTNG 122
                 || :   :  :||   :  | |  |::     :   |  :|| -   ::|   |
Db        305 VAVNGQLKANAGANAAVIGTNNGAGRAAGFVVSVDNGKVATIDGQVYAKDMVIQSANAVG 364

Qy        123 DTT----VHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKD---VLNAGWNIK 175
                 | :   |:|   |  :  | |  :|   ||     |  :        ||  :
Db        365 QVNFRHIVDVGTDGTTAFKTAASKVAITQNSNFGTTDFGNLAAQIIVPNTMTLNGNFTGD 424

Qy        176 GVKPGTTASDNVDFVRTYD---TVEFLSADTKTTTVN----VESKDNG-------KKTEV 221
                || ||      | |:|   |:   |||       |    :|:  |           |:
Db        425 ASNPGNTAG-----VITFDANGTLASASADANVAVTNNITAIEASGAGVVQLSGTHAAEL 479

Qy        222 KIGAKTSVIKEKDGKLVTGKDKG--------ENGSSTDEGEGLVTAKEVIDAVNKAGWRM 273
              ::|   || |  || :: ||              |: | :|   :|     |  |  |
Db        480 RLGNAGSVFKLADGTVINGKVNQTALVGGALAAGTITLDGSATITG----DIGNAGG--- 532

Qy        274 KTTTANGQTGQADKFETVT---------SGTNVTFASGKGTTATVSKDDQGNITVMYDV- 323
                | |   |  :|:|             :|  : | :  ||    |   | || | :|:
Db        533 -AAALQGITLANDATKTLTLGGANIIGANGGTINFQANGGTIKLTS--TQNNIVVDFDLA 589

Qy        324 ------NVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKM-DETVNIN---A 373
                   | || ::   |    |     |  ::  :   |:  ||  : |  | ||
Db        590 IATDQTGVVDASSLTNAQTLTINGKIGTVGANNKTLGQFNIGSSKTVLSDGDVAINELVI 649

Qy        374 GNN--IEITRNGKNIDIATSMTPQ----FSSV-----SLGAGADAPTLSVDGDALNVGSK 422
              |||  ::   |   |   |:   |    |: |      :|  | :  : :      :| |||
Db        650 GNNGAVQFAHNTYLITRTTNAAGQGKIIFNPVVNNNTTLATGTNLGSATNPLAEINFGSK 709

Qy        423 KDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDG-NARAGIAQAIATAGLVQ 481
                : ||  ||     ||:        |: : | |  | |  :|      | |  |
Db        710 GAANVDTVLNVGKGVNL-YATNITTTDA---NVGSFIFNAGGTNIVSG------TVGGQQ 759

Qy        482 -------AYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRG 528
                    |   | ::  :|  |: |    |   |::  |||:     || :  |
Db        760 GNKFNTVALDNGTTVKFLGNATFNGNTTIAAN-STLQIGGNYTADFVASADGTG 812


RESULT 12
YPJA_ECOLI
ID   YPJA_ECOLI     STANDARD;      PRT;  1569 AA.
AC   P52143; P76610; P77017; P77019;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Hypothetical outer membrane protein ypjA.
GN   YPJA OR B2647.
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12 / MG1655;
RX   MEDLINE=97426617; PubMed=9278503;
RA   Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA   Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA   Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA   Mau B., Shao Y.;
RT   "The complete genome sequence of Escherichia coli K-12.";
RL   Science 277:1453-1474(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12;
RX   MEDLINE=97349980; PubMed=9205837;
RA   Yamamoto Y., Aiba H., Baba T., Hayashi K., Inada T., Isono K.,
RA   Itoh T., Kimura S., Kitagawa M., Makino K., Miki T., Mitsuhashi N.,
RA   Mizobuchi K., Mori H., Nakade S., Nakamura Y., Nashimoto H.,
RA   Oshima T., Oyama S., Saito N., Sampei G., Satoh Y., Sivasundaram S.,
RA   Tagami H., Takahashi H., Takeda J., Takemoto K., Uehara K., Wada C.,
RA   Yamagata S., Horiuchi T.;
RT   "Construction of a contiguous 874-kb sequence of the Escherichia coli
RT   - K12 genome corresponding to 50.0-68.8 min on the linkage map and
RT   analysis of its sequence features.";
RL   DNA Res. 4:91-113(1997).
CC   -!- SUBCELLULAR LOCATION: Outer membrane (Potential).
CC   -!- SIMILARITY: STRONG, TO BORDETELLA PERTACTIN.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U36840; AAA79815.1; ALT_SEQ.
DR   EMBL; AE000350; AAC75695.1; -.
DR   EMBL; D90889; BAA16514.1; ALT_INIT.
DR   EMBL; D90890; BAA16518.1; ALT_INIT.
DR   PIR; A65044; A65044.
DR   EcoGene; EG13213; ypjA.
DR   InterPro; IPR006315; Autotransport.
DR   InterPro; IPR005546; Autotransporter.
DR   InterPro; IPR004899; Pertactin.
DR   InterPro; IPR003991; Pertactin_C.
DR   Pfam; PF03797; Autotransporter; 1.
DR   Pfam; PF03212; Pertactin; 1.
DR   PRINTS; PR01484; PRTACTNFAMLY.
DR   TIGRFAMs; TIGR01414; autotrans_barl; 2.
KW   Hypothetical protein; Outer membrane; Complete proteome.
SQ   SEQUENCE   1569 AA;  162774 MW;  0D006A9A32154596 CRC64;

  Query Match             6.4%;  Score 176.5;  DB 1;  Length 1569;
  Best Local Similarity   19.6%;  Pred. No. 0.13;
  Matches  104;  Conservative   94;  Mismatches  255;  Indels   77;  Gaps   19;

Qy         29 GAGEKEKVEENSDWAVYFNEKGVLTAREITLKAGDNLKIKQNGTNFTYSLKKDLTDLTSV 88
              |:  |  :::    :|  | |   |     :  |  : :  | ||          : :
Db        373 GSANKTIIKDKGTMSVLTNAKADAT----RIDNGGVMDVAGNATN----------TIING 418

Qy         89 GTEKLSFSAHGNKVNITSDTKGLNFAKETAGTNGDTTVHLNG----------IGSTLT- 136
              ||: ::         || | |: :      :|   |||:  :|          ||| ::
Db        419 GTQNINNYGIATGTNINSGTQNIK-----SGGKADTTIISSGSRQVVEKDGTAIGSNISA 473

Qy        137 --DTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKG---VKPGTTASDNVDFVR 191
                 :: ||   :  |      ::  :| | :   || :|:|   :   |      ::|
Db        474 GGSLIVYTGGIAHGVN------QETGSALVANT-GAGTDIEGYNKLSHFTITGGEANYVV 526

Qy        192 TYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKDKGENGSSTD 251
              :| | |:   ||:  |  : | | |  |:  |||   :  :| ::  :| ||
Db        527 LENTGE-LTVVAKTSAKNT-TIDTGGKLIVQKEAKTDSTRLNNGGVLEVQDGGEAKHVEQ 584

Qy        252 EGEGLV----TAKEVIDAVNKAG--WRMKTTTANGQTGQADKFETVTSGTNV--TFASGK 303
              :  | :    |:  :|:  |  |  : :: : |     :     || :|:    |  :
Db        585 QSGGALIASTTSGTLIEGTNSYGDAFYIRNSEAKNVVLENAGSLTVVTGSRAVDTIINAN 644

Qy        304 GTTATVSKD-----DQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNV 358
              |       ||      :      :|      |  |:   : : : | ::|   |  ::: |
Db        645 GKMDVYGKDVGTVLNSAGTQTIYASATSDKANIKGGKQTVYGLATEANIESGEQIVDGGS 704

Qy        359 SPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGDALN 418
              :       |  :|| |    :   || |:  | :      :     |:   ::     :|
Db        705 T-------EKTHINGGTQ-TVQNYGKAIN--TDIVSGLQQIMANGTAEGSIINGGSQVVN 754

Qy        419 VGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAG 478
              |   :|  :        ::|  | : |     |    |  | |    |:| :|
Db        755 EGGLAENSVLNDGGTLDVREKGSATGIQQSSQGALVATTRATRVTGTRADGVAFSIEQGA 814
```

```
Qy        479 LVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWII--KGTASGNS 526
                  :::   ||    |:  :   : :: ||  |:  | ||:| :
Db        815 -------ANNILLANGGVLTVESDTSSDKTQVNMGGREIVKTKATATGTT 857


RESULT 13
ICEN_XANCT
ID   ICEN_XANCT     STANDARD;      PRT;  1567 AA.
AC   P18127;
DT   01-NOV-1990 (Rel. 16, Created)
DT   01-NOV-1990 (Rel. 16, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Ice nucleation protein.
GN   INAX.
OS   Xanthomonas campestris (pv. translucens).
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC   Xanthomonadaceae; Xanthomonas.
OX   NCBI_TaxID=343;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=X56S;
RX   MEDLINE=91080859; PubMed=2259339;
RA   Zhao J., Orser C.S.;
RT   "Conserved repetition in the ice nucleation gene inaX from
RT   Xanthomonas campestris pv. translucens.";
RL   Mol. Gen. Genet. 223:163-166(1990).
CC   -!- FUNCTION: ICE NUCLEATION PROTEINS ENABLE BACTERIA TO NUCLEATE
CC       CRYSTALLIZATION IN SUPERCOOLED WATER.
CC   -!- SUBCELLULAR LOCATION: Outer membrane (By similarity).
CC   -!- DOMAIN: CONTAINS 153 IMPERFECT REPEATS OF THE CONSENSUS
CC       OCTAPEPTIDE A-G-Y-G-S-T-L-T; FURTHER ON A 16-RESIDUE AND A
CC       REGIONAL 48-RESIDUE PERIODICITY IS SUPERIMPOSED.
CC   -!- MISCELLANEOUS: A STRUCTURAL MODEL IS SUGGESTED IN WHICH THE ICE
CC       NUCLEATION PROTEIN DISPLAYS A SYMMETRY RELATED TO THAT OF ICE.
CC   -!- SIMILARITY: BELONGS TO THE BACTERIAL ICE NUCLEATION PROTEIN
CC       FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X52970; CAA37140.1; -.
DR   HSSP; P06620; 1INA.
DR   InterPro; IPR000258; Ice_nucleatn.
DR   Pfam; PF00818; Ice_nucleation; 81.
DR   PRINTS; PR00327; ICENUCLEATN.
DR   PROSITE; PS00314; ICE_NUCLEATION; 57.
KW   Ice nucleation; Repeat; Outer membrane.
SQ   SEQUENCE   1567 AA;  152548 MW;  C8B451D959ECAD63 CRC64;

  Query Match             6.3%;  Score 174;  DB 1;  Length 1567;
  Best Local Similarity   21.5%;  Pred. No. 0.18;
  Matches  120;  Conservative   69;  Mismatches  246;  Indels  122;  Gaps   23;

Qy         29 GAGEKEKVEENSDWAVYFNEKGVLTAREITLKAG---------DNLKIKQNGTNFTYSLK 79
              | |  :   | ||    :   |   | : || ||         |:      |:  |
Db        605 GYGSTQTAREGSDVTAGYGSTGTAGA-DSTLIAGYGSTQTSGSDSSLTAGYGSTQTARKG 663

Qy         80 KDLT----DLTSVGTEKLSFSAHGNKVNITSD---TKGLNFAKETA----------G 119
                |:|       : | :    : :|:     ||   | |   : :||          |
Db        664 SDVTAGYGSTGTAGADSTLIAGYGSTQTSGSDSSLTAGYG-STQTARKGSDVTAGYGSTG 722

Qy        120 TNGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKP 179
              | |  :  : | |||  |     :|: :::|    :    :: : |     ||:   |
Db        723 TAGADSTLIAGYGSTQT-----SGSDSSLTAGYGSTQTARKGSDV----TAGYGSTG--- 770

Qy        180 GTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVT 239
                 ||  :   :  | : :   :|:  |     ::   | :::  |  ::     |  |:
Db        771 --TAGADSTLIAGYGSTQTSGSDSSLTAGYGSTQTARKGSDITAGYGSTGTAGADSTLIA 828

Qy        240 GKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTF 299
              |      ||:   |          |:   ||:    |   |    |    | |:| : |
Db        829 G-----YGSTQTSGS---------DSSLTAGYGSTQTAREGSDVTAGYGSTGTAGADSTL 874

Qy        300 ASGKGTTATVSKDDQGNITVMY----------DVNVGDALNVNQLQNSGWNLDSKAVAGS 349
                || |:| |   |   ::|  |          ||  |                |   ||:
Db        875 ISGYGSTQTAGSD--SSLTAGYGSTQTARKGSDVTAG--------------YGSTGTAGA 918

Qy        350 SGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPT 409
                 :|:|  |      | ::    |:  :  | |      : ||  : |    ||||:
Db        919 DSTLIAGYGSTQTSGSDSSLTAGYGST-QTARKG------SDMTAGYGSTGT-AGADSTL 970

Qy        410 LSVDGDALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAG 469
              ::  |     ||         :::  |             |  |      |   |    |
Db        971 IAGYGSTQTSGSD--------SSLTAGY------------GSTQTAREGSDVTAGYGSTG 1010

Qy        470 IAQA----IATAGLVQAYLPGKSMMAIGGGTYRGEAGYAI--GYSSISDGG--NWIIKGT 521
              | |    ||  |  |      |: |  | |     |  :  || |    |  : :| |
Db       1011 TAGADSTLIAGYGSTQTAGSDSSLTAGYGSTQTARQGSDVTAGYGSTGTAGADSTLIAGY 1070

Qy        522 ASGNSRGHFGASASVGY 538
              |  : |   :| : ||
Db       1071 GSTQTAGS-DSSLTAGY 1086


RESULT 14
WAPA_BACSU
ID   WAPA_BACSU     STANDARD;      PRT;  2334 AA.
AC   Q07833;
DT   01-OCT-1994 (Rel. 30, Created)
DT   01-OCT-1994 (Rel. 30, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Wall-associated protein precursor.
GN   WAPA OR N17G.
OS   Bacillus subtilis.
OC   Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX   NCBI_TaxID=1423;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=168;
RX   MEDLINE=93302506; PubMed=8316082;
RA   Foster S.J.;
RT   "Molecular analysis of three major wall-associated proteins of
RT   Bacillus subtilis 168: evidence for processing of the product of a
RT   gene encoding a 258 kDa precursor two-domain ligand-binding
RT   protein.";
RL   Mol. Microbiol. 8:299-310(1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=168 / BGSC1A1;
RX   MEDLINE=95219088; PubMed=7704263;
RA   Yoshida K.-I., Sano H., Seki S., Oda M., Fujimura M., Fujita Y.;
RT   "Cloning and sequencing of a 29 kb region of the Bacillus subtilis
RT   genome containing the hut and wapA loci.";
RL   Microbiology 141:337-343(1995).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=168 / BGSC1A1;
RX   MEDLINE=97124196; PubMed=8969509;
RA   Yoshida K.-I., Shindo K., Sano H., Seki S., Fujimura M., Yanai N.,
RA   Miwa Y., Fujita Y.;
RT   "Sequencing of a 65 kb region of the Bacillus subtilis genome
RT   containing the lic and cel loci, and creation of a 177 kb contig
RT   covering the gnt-sacXY region.";
RL   Microbiology 142:3113-3123(1996).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=168;
RX   MEDLINE=98044033; PubMed=9384377;
RA   Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA   Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA   Borriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
```

```
RA   Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA   Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA   Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
RA   Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA   Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA   Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
RA   Guiseppi G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,
RA   Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA   Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA   Kobayashi Y., Koetter P., Koningstein G., Krogh S., Kumano M.,
RA   Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA   Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
RA   Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA   Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA   Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA   Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA   Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,
RA   Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA   Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
RA   Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA   Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA   Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA   Viari A., Wambutt R., Wedler E., Wedler H., Weitzenegger T.,
RA   Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA   Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;
RT   "The complete genome sequence of the Gram-positive bacterium Bacillus
RT   subtilis.";
RL   Nature 390:249-256(1997).
CC   -!- FUNCTION: STILL UNKNOWN. NOT INVOLVED IN CELL MEMBRANE METABOLISM,
CC       MOTILITY, SECRETION OR DIFFERENTIATION.
CC   -!- SUBCELLULAR LOCATION: MEMBRANE-ASSOCIATED. MAY BE RELEASED
CC       INTO THE MEDIUM.
CC   -!- DOMAIN: HAS TWO LIGAND-BINDING DOMAINS; THE N-TERMINUS, HAS THREE
CC       101 AA REPEATS WHICH ARE RESPONSIBLE FOR CELL WALL BINDING; THE
CC       C-TERMINUS CONSISTS OF TWO BLOCKS OF RESIDUES WITH A CONSERVED
CC       MOTIF REPEATED 31 TIMES.
CC   -!- SIMILARITY: THE REPEATED MOTIF IN THE C-TERMINUS HAS SOME
CC       SIMILARITY TO THE REPEAT IN E.COLI RHS GROUP OF PROTEINS (RHSA-D).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L05634; AAA22883.1; -.
DR   EMBL; D31856; BAA06656.1; -.
DR   EMBL; D29985; BAA06260.1; -.
DR   EMBL; D83026; BAA11683.1; -.
DR   EMBL; Z99124; CAB15959.1; -.
DR   PIR; S32920; S32920.
DR   SubtiList; BG10797; wapA.
DR   InterPro; IPR003305; CBM_CenC.
DR   InterPro; IPR006530; YD.
DR   Pfam; PF02018; CBM_4_9; 1.
DR   TIGRFAMs; TIGR01643; YD_repeat_2x; 17.
KW   Cell wall; Repeat; Signal; Complete proteome.
FT   SIGNAL       1     28       OR 32 (POTENTIAL).
FT   CHAIN       29   2334       WALL-ASSOCIATED PROTEIN.
FT   DOMAIN     504    869       3 X 101 AA APPROXIMATE TANDEM REPEATS.
FT   REPEAT     504    605       1-1.
FT   REPEAT     636    736       1-2.
FT   REPEAT     769    869       1-3.
FT   DOMAIN    1021   2139       31 X 21 AA APPROXIMATE TANDEM REPEATS OF
FT                               X(4)-G-X(4)-[YF]-X-D-X(2)-G-X(4).
FT   REPEAT    1021   1040       2-1.
FT   REPEAT    1042   1061       2-2.
FT   REPEAT    1063   1082       2-3.
FT   REPEAT    1083   1102       2-4.
FT   REPEAT    1109   1128       2-5.
FT   REPEAT    1129   1148       2-6.
FT   REPEAT    1150   1169       2-7.
FT   REPEAT    1174   1193       2-8.
FT   REPEAT    1199   1218       2-9.
FT   REPEAT    1219   1238       2-10.
FT   REPEAT    1646   1665       2-11.
FT   REPEAT    1667   1686       2-12.
FT   REPEAT    1690   1709       2-13.
FT   REPEAT    1711   1730       2-14.
FT   REPEAT    1732   1751       2-15.
FT   REPEAT    1753   1772       2-16.
FT   REPEAT    1795   1814       2-17.
FT   REPEAT    1820   1839       2-18.
FT   REPEAT    1840   1859       2-19.
FT   REPEAT    1861   1880       2-20.
FT   REPEAT    1887   1906       2-21.
FT   REPEAT    1908   1927       2-22.
FT   REPEAT    1929   1948       2-23.
FT   REPEAT    1969   1982       2-24 (APPROXIMATE).
FT   REPEAT    1983   2002       2-25.
FT   REPEAT    2008   2027       2-26.
FT   REPEAT    2028   2047       2-27.
FT   REPEAT    2051   2070       2-28.
FT   REPEAT    2071   2090       2-29.
FT   REPEAT    2093   2112       2-30.
FT   REPEAT    2120   2139       2-31.
SQ   SEQUENCE   2334 AA;  258329 MW;  B75138CCD278BAA3 CRC64;

  Query Match             6.3%;  Score 173.5;  DB 1;  Length 2334;
  Best Local Similarity   23.1%;  Pred. No. 0.3;
  Matches  110;  Conservative   63;  Mismatches  169;  Indels  135;  Gaps   23;

Qy         33 KEKVEENSDWAVYFNEKGVLTAREITLKAGDNLKIKQNGTNFTYSLKKDLTDLTSVGTEK 92
              | | : |:    |||:||    : :    | |     | | :||: |  || :|     |
Db       1001 KTKDQTNA----YFNKKGGKLQKVVD---GHN-----NATVYTYNDKNQLTAITDASGRK 1048

Qy         93 LSFS--AHGNKVNITS-DTKGLNFAKETAGTNGDTTVHLNGIGSTLTDTLLNTGATTNVT 149
              |:|:   :|:  :||    | : :: |           | :   :|||    |  |:
Db       1049 LTFTYDENGHVTSITGPKNKKVTYSYE-----------NDLLKKVTDT---DGTVTSYD 1093

Qy        150 NDNVTDDEKKRAASVKDVLNAGWNIKGVKPGTTA-------------SDNVDFVRTYDTV 196
              |:     |: :|         |   || |                 :   :| :||
Db       1094 YDSEGRLVKQYSA----------NSTEAKPVFTEYQYSGHRLEKAINAKKETYVYSYD-- 1141

Qy        197 EFLSADTKTTTVNVESKDNGKKTEV---KIGAKTSVIKEKDGKLVT------GKDKGENG 247
                || ||  :   :: ||:| :    : |    || : :|  :|      | :  |:
Db       1142 ----ADKKTLLM---TQPNGRKVQYGYNEAGNPIQVIDDAEGLKITTNTKYEGNNVVEDV 1194

Qy        248 SSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTS-----GTNVTFASG 302
              | | |  |                    :: | ||   |||     ||  |:
Db       1195 DPNDVGTGKAT---------------------ESYQYDKDGNVTSVKDAYGTE-TYEYN 1231

Qy        303 KGTTATVSKDDQGNIT-VMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPS 361
              |    |  || :||:| : ||                ||: :    |||  |  |
Db       1232 KNNDVTKMKDTEGNVTDIAYD----------------GLDAVSETDQSGKSSSAAVYDK 1274

Qy        362 KG-KMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGDALNVG 420
              | ::  : :::|  |  | ::|       |   | |  :| |  |   :|| |   |
Db       1275 YGNQIQSSKDLSASTN--ILKDG-------SFEAQKSGWNLTASKDRRKISVIADKSGVL 1325

Qy        421 SKKDNKPVRITNVAPGVKEG--DVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIA 475
              |      |   : : |   |    |   :|:       |    : | :  :|:: |
Db       1326 SGSKALEVLSQSTSAGTDHGYSSATQTVELE------PNTTYTLSGKIKTDLAKSRA 1376


RESULT 15
SLAP_CAMFE
ID   SLAP_CAMFE     STANDARD;      PRT;   933 AA.
AC   P35827;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   01-OCT-1994 (Rel. 30, Last annotation update)
DE   S-layer protein (Surface array protein) (SAP).
GN   SAPA.
```

```
OS   Campylobacter fetus.
OC   Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
OC   Campylobacteraceae; Campylobacter.
OX   NCBI_TaxID=196;
RN   [1]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 1-20.
RC   STRAIN=84-32 / 23D;
RX   MEDLINE=90354448; PubMed=2387868;
RA   Blaser M.J., Gotschlich E.C.;
RT   "Surface array protein of Campylobacter fetus. Cloning and gene
RT   structure.";
RL   J. Biol. Chem. 265:14529-14535(1990).
RN   [2]
RP   ERRATUM.
RX   MEDLINE=91035477; PubMed=2229082;
RA   Blaser M.J., Gotschlich E.C.;
RL   J. Biol. Chem. 265:19372-19372(1990).
CC   -!- FUNCTION: THE S-LAYER IS A PARACRYSTALLINE MONO-LAYERED ASSEMBLY
CC       OF PROTEINS WHICH COAT THE SURFACE OF BACTERIA. THIS PROTEIN IS
CC       CRITICAL FOR VIRULENCE.
CC   -!- SUBCELLULAR LOCATION: CELL WALL. THIS BACTERIUM IS COVERED BY A
CC       S-LAYER WITH HEXAGONAL SYMMETRY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; J05577; AAA23032.1; -.
KW   Cell wall; S-layer.
SQ   SEQUENCE   933 AA;  96757 MW;  F88C729B4BA5B1E9 CRC64;

  Query Match             6.3%;  Score 173;  DB 1;  Length 933;
  Best Local Similarity   22.4%;  Pred. No. 0.11;
  Matches  138;  Conservative   72;  Mismatches  243;  Indels  164;  Gaps   31;

Qy         38 ENSDWAVYFNE--KGVLTA--REITLKAGDNLKIKQNGTNF-TYSLKKDLT---DLTSVG 89
              : : |:|   :    : ||  |   |     | | :  ||    ::  ||    |    |
Db        234 DKTRWSVLGRDTLNAIFTAITRAALLTDQAELIITKRRTNVENINIISDLETSGDFVFNG 293

Qy         90 TEKLSFSAHGNKVNITSD-TKGLNFAKETAGT-----------------------NGDT 124
               ||: |:  |: |:  :| :| :|   || ||                       :  |
Db        294 YEKVGFNVLGDIVSFATDASKSVNV--ETTGTITAFTAAGTGKVDVVAGKISALTADSRT 351

Qy        125 TVHLNGIGSTLTDTLLNTGATTNVTNDNVTD-----------DEKKRAA---------- 162
              :|:|     |:| |  |   : |:      |            : : | |
Db        352 SVNLTATNDTITLTSANAATSVNLKQRQAKDATITSAMQQKYNNRRNRIATITSATAVEN 411

Qy        163 -SVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVN-VESKDNG---- 216
              :||   |   |    |  |   ||       |           :|:| : |  ||
Db        412 LTVKHATNVALNGGMDKLATVTLDNAALTAAIDI-------KSASTLNLINSSVNGPKHL 464

Qy        217 ----------KKTEVKIGAKTSVIKEK---------DGKLVTGKDKGENGSSTDEGEG-- 255
                       |:   |:   |:   ::          |  |        :  | |  | |
Db        465 YSSKRRYCKFKRAAAKVKLNTTAATDQTVTLKANATDNSLEFDSATAKTTSVTASGSGKT 524

Qy        256 -LVTAKEVIDAVNKAGWRMKTTTAN-------GQTGQADKFETVTSGTNVTFASGKGTTA 307
               ::   ||   ||      : ||  |        |:|||  || :| :||         |||
Db        525 LVIKGAEVETLVN-----IDTTAFNALQSVSFGKTGQGGKF-SVKTGTGDDKIEFVGTTL 578

Qy        308 TVSK--DDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAG---SSGKVISGNVSPSK 362
              |      |  || |:         : |   ::|      | |||    ||    :  :  :|
Db        579 TEGSVIDAPGNDTIAMKSAALTSANFTMIKNIENVAISDAVATADLSSSAFKNSVIITTK 638

Qy        363 GKMDETVNINAGNNIEITRNGKN-----------------IDIATSMTPQFSSVSLGAG 404
                 | |: ||    |  |                       : |      : ::::||
Db        639 EAADTTLTINKDQVINFTAADAGSVKLITVKLNDVTALMIVVKIVLDAAAKDTNIALGTA 698

Qy        405 ADAPTLSVDG--DALNVGS-KKDNKPVRITNVAPGVKEGDVTNV----AQL----KGVAQ 453
              |    | :|   : ||: |  |   |    |     |  |||::     |:     | |
Db        699 AADKALVIDTGIETLNITSLVKATSPETTANTV-NAKLTDVTSIIIDGMQITLGHAGTAG 757

Qy        454 NLNNRIDNVDGNA-RAGI---AQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSS 509
                  :::  :| :| :||:    | || | |         :   | ||:         |  |
Db        758 TDYSKVSMIDASALKAGLTFDASAI-TLG---------ANATIKGGS---------GADS 798

Qy        510 IS-DGGNWIIKGTASGN 525
              |:  ||| ::   | |:
Db        799 ITVKGGNIVVDLVAGGD 815


Search completed: October  6, 2003, 09:24:09
Job time : 11.339 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: October 6, 2003, 09:13:10 ; Search time 39.661 Seconds
(without alignments)
3513.485 Million cell updates/sec

Title: US-09-771-382-33
Perfect score: 2757
Sequence: 1 NNEEQEEYLYLHPVQRTVAV.....TASGNSRGHFGASASVGYQW 540

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_23:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2727	98.9	591	2	Q9JPS7	Q9jps7 neisseria m
2	2727	98.9	591	16	Q9JR18	Q9jr18 neisseria m
3	2720	98.7	591	2	Q93QY3	Q93qy3 neisseria m
4	2695.5	97.8	592	2	Q9AQF0	Q9aqf0 neisseria m
5	2624.5	95.2	590	2	Q9JPS3	Q9jps3 neisseria m
6	2512	91.1	594	2	Q93QY4	Q93qy4 neisseria m
7	2511	91.1	600	2	Q9JPS6	Q9jps6 neisseria m
8	2501	90.7	594	2	Q9JPI3	Q9jpi3 neisseria m
9	2501	90.7	594	2	Q9JPS2	Q9jps2 neisseria m
10	2493	90.4	594	2	Q9JPH7	Q9jph7 neisseria m
11	2491	90.4	598	2	Q9JPR9	Q9jpr9 neisseria m
12	2478	89.9	598	2	Q9JPS0	Q9jps0 neisseria m
13	2478	89.9	598	2	Q93QY5	Q93qy5 neisseria m
14	2478	89.9	598	2	Q9JPT0	Q9jpt0 neisseria m
15	2442.5	88.6	526	2	Q9JPS4	Q9jps4 neisseria m
16	2440.5	88.5	530	2	Q9JPS1	Q9jps1 neisseria m
17	2430.5	88.2	599	2	Q9JPR8	Q9jpr8 neisseria m
18	2371	86.0	599	2	Q9JPS8	Q9jps8 neisseria m
19	2364	85.7	592	2	Q9JPS9	Q9jps9 neisseria m
20	2328	84.4	598	2	Q9JPR7	Q9jpr7 neisseria m
21	2321	84.2	592	2	Q93QY2	Q93qy2 neisseria m
22	2284.5	82.9	595	2	Q9JPH0	Q9jph0 neisseria m
23	2284.5	82.9	600	2	Q9JPS5	Q9jps5 neisseria m
24	2252.5	81.7	589	2	Q9JPI0	Q9jpi0 neisseria m
25	2241.5	81.3	589	2	Q93QY1	Q93qy1 neisseria m
26	2212	80.2	592	16	Q9JQW4	Q9jqw4 neisseria m
27	1031.5	37.4	2353	2	P71401	P71401 haemophilus
28	915	33.2	1210	2	Q8GM74	Q8gm74 haemophilus
29	908	32.9	1098	2	Q48152	Q48152 haemophilus
30	907	32.9	1210	2	Q8GM75	Q8gm75 haemophilus
31	904.5	32.8	1204	2	Q8GM76	Q8gm76 haemophilus
32	904	32.8	1096	2	Q8GM79	Q8gm79 haemophilus
33	591	21.4	1002	2	Q8GM78	Q8gm78 haemophilus
34	588.5	21.3	1004	2	Q8GM77	Q8gm77 haemophilus
35	429	15.6	1299	16	Q9F3X6	Q9f3x6 pasteurella
36	393	14.3	2059	16	Q9PD50	Q9pd50 xylella fas
37	385	14.0	2314	2	Q8KQM8	Q8kqm8 moraxella c
38	374	13.6	1964	2	Q8KQM9	Q8kqm9 moraxella c
39	371.5	13.5	1461	16	Q8ZL64	Q8zl64 salmonella
40	369.5	13.4	1107	16	Q9F2D8	Q9f2d8 salmonella
41	365.5	13.3	1190	16	Q9PC04	Q9pc04 xylella fas
42	364	13.2	1778	16	Q8FCB2	Q8fcb2 escherichia
43	361.5	13.1	2712	16	Q9F3X5	Q9f3x5 pasteurella
44	358	13.0	1588	16	Q8XDG4	Q8xdg4 escherichia
45	345	12.5	688	2	Q8RQ60	Q8rq60 actinobacil

ALIGNMENTS

RESULT 1
Q9JPS7
ID Q9JPS7 PRELIMINARY; PRT; 591 AA.
AC Q9JPS7;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Outer membrane protein GNA992.
GN GNA992.
OS Neisseria meningitidis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=487;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BZ147;
RX MEDLINE=20175756; PubMed=10710308;
RA Pizza M., Scarlato V., Masignani V., Giuliani M.M., Arico' B.,
RA Comanducci M., Jennings G.T., Baldi L., Bartolini E., Capecchi B.,
RA Galeotti C.L., Luzzi E., Manetti R., Marchetti E., Mora M., Nuti S.,
RA Ratti G., Santini L., Savino S., Scarselli M., Storni E., Zuo P.,
RA Broeker M., Hundt E., Knapp B., Blair E., Mason T., Tettelin H.,
RA Hood D.W., Jeffries A.C., Saunders N.J., Granoff D.M., Venter C.,
RA Moxon E.R., Grandi G., Rappuoli R.;
RT "Identification of Vaccine Candidates Against Serogroup B
RT Meningococcus by Whole-Genome Sequencing.";
RL Science 287:1816-1820(2000).
DR EMBL; AF226366; AAF42515.1; -.
DR InterPro; IPR005594; YadA.
DR Pfam; PF03895; YadA; 1.
SQ SEQUENCE 591 AA; 62113 MW; 533453CAE5A91E1F CRC64;

Query Match 98.9%; Score 2727; DB 2; Length 591;
Best Local Similarity 99.3%; Pred. No. 1.2e-107;
Matches 536; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

```
Qy          1 NNEEQEEYLYLHPVQRTVAVLIVNSDKEGAGEKEKVEENSDWAVYFNEKGVLTAREITLK 60
              ||||||| ||| ||||||||||||||| |||||||||||||||||||||||||||||||
Db         52 NNEEQEEDLYLDPVQRTVAVLIVNSDKEGTGEKEKVEENSDWAVYFNEKGVLTAREITLK 111
```

```
Qy          61 AGDNLKIKQNGTNFTYSLKKDLTDLTSVGTEKLSFSAHGNKVNITSDTKGLNFAKETAGT 120
               ||||||||||||||||||||||||||||||||||||:||||||||||||||||||||||
Db         112 AGDNLKIKQNGTNFTYSLKKDLTDLTSVGTEKLSFSANGNKVNITSDTKGLNFAKETAGT 171

Qy         121 NGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPG 180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         172 NGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPG 231

Qy         181 TTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTG 240
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         232 TTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTG 291

Qy         241 KDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFA 300
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         292 KDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFA 351

Qy         301 SGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSP 360
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         352 SGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSP 411

Qy         361 SKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGDALNVG 420
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         412 SKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGDALNVG 471

Qy         421 SKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLV 480
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         472 SKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLV 531

Qy         481 QAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 540
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         532 QAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 591
```

```
RESULT 2
Q9JR18
ID   Q9JR18      PRELIMINARY;      PRT;   591 AA.
AC   Q9JR18;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Outer membrane protein GNA992 (Adhesin) (NhhA outer membrane
DE   protein).
GN   GNA992 OR NMB0992 OR NHHA.
OS   Neisseria meningitidis, and
OS   Neisseria meningitidis (serogroup B).
OC   Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC   Neisseriaceae; Neisseria.
OX   NCBI_TaxID=487, 491;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MC58 / Serogroup B, BZ169, BZ83, and H44/76;
RX   MEDLINE=20175756; PubMed=10710308;
RA   Pizza M., Scarlato V., Masignani V., Giuliani M.M., Arico' B.,
RA   Comanducci M., Jennings G.T., Baldi L., Bartolini E., Capecchi B.,
RA   Galeotti C.L., Luzzi E., Manetti R., Marchetti E., Mora M., Nuti S.,
RA   Ratti G., Santini L., Savino S., Scarselli M., Storni E., Zuo P.,
RA   Broeker M., Hundt E., Knapp B., Blair E., Mason T., Tettelin H.,
RA   Hood D.W., Jeffries A.C., Saunders N.J., Granoff D.M., Venter C.,
RA   Moxon E.R., Grandi G., Rappuoli R.;
RT   "Identification of Vaccine Candidates Against Serogroup B
RT   Meningococcus by Whole-Genome Sequencing.";
RL   Science 287:1816-1820(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MC58 / Serogroup B;
RX   MEDLINE=20175755; PubMed=10710307;
RA   Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
RA   Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
RA   Nelson W.C., Gwinn M.L., DeBoy R., Peterson J.D., Hickey E.K.,
RA   Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
RA   Mason T., Ciecko A., Parksey D.S., Blair E., Cittone H., Clark E.B.,
RA   Cotton M.D., Utterback T.R., Khouri H., Qin H., Vamathevan J.,
RA   Gill J., Scarlato V., Masignani V., Pizza M., Grandi G., Sun L.,
RA   Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;
RT   "Complete genome sequence of Neisseria meningitidis serogroup B strain
RT   MC58.";
RL   Science 287:1809-1815(2000).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   SPECIES=N.meningitidis; STRAIN=PMC21;
RA   Peak I.R., Srikhanta Y., Dieckelman M., Moxon R., Jennings M.P.;
RT   "Identification and characterization of a gene encoding a novel outer
RT   membrane protein of Neisseria meningitidis.";
RL   Submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF226375; AAF42524.1; -.
DR   EMBL; AE002450; AAF41395.1; -.
DR   EMBL; AF226367; AAF42516.1; -.
DR   EMBL; AF226370; AAF42519.1; -.
DR   EMBL; AF226374; AAF42523.1; -.
DR   EMBL; AF157611; AAK68872.1; -.
DR   TIGR; NMB0992; -.
DR   InterPro; IPR005594; YadA.
DR   Pfam; PF03895; YadA; 1.
KW   Complete proteome.
SQ   SEQUENCE   591 AA;  62112 MW;  7C22F3CAE7F73EC6 CRC64;

  Query Match            98.9%;  Score 2727;  DB 16;  Length 591;
  Best Local Similarity  99.3%;  Pred. No. 1.2e-107;
  Matches  536;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy           1 NNEEQEEYLYLHPVQRTVAVLIVNSDKEGAGEKEKVEENSDWAVYFNEKGVLTAREITLK 60
               ||||||| ||| ||||||||||||||| |||||||||||||||||||||||||||||||
Db          52 NNEEQEEDLYLDPVQRTVAVLIVNSDKEGTGEKEKVEENSDWAVYFNEKGVLTAREITLK 111

Qy          61 AGDNLKIKQNGTNFTYSLKKDLTDLTSVGTEKLSFSAHGNKVNITSDTKGLNFAKETAGT 120
               ||||||||||||||||||||||||||||||||||||:||||||||||||||||||||||
Db         112 AGDNLKIKQNGTNFTYSLKKDLTDLTSVGTEKLSFSANGNKVNITSDTKGLNFAKETAGT 171

Qy         121 NGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPG 180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         172 NGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPG 231

Qy         181 TTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTG 240
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         232 TTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTG 291

Qy         241 KDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFA 300
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         292 KDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFA 351

Qy         301 SGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSP 360
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         352 SGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSP 411

Qy         361 SKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGDALNVG 420
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         412 SKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGDALNVG 471

Qy         421 SKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLV 480
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         472 SKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLV 531

Qy         481 QAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 540
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         532 QAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 591
```

```
RESULT 3
Q93QY3
ID   Q93QY3      PRELIMINARY;      PRT;   591 AA.
AC   Q93QY3;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   NhhA outer membrane protein.
```

```
GN   NHHA.
OS   Neisseria meningitidis.
OC   Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC   Neisseriaceae; Neisseria.
OX   NCBI_TaxID=487;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=EG329;
RA   Peak I.R., Srikhanta Y., Dieckelman M., Moxon R., Jennings M.P.;
RT   "Identification and characterization of a gene encoding a novel outer
RT   membrane protein of Neisseria meningitidis.";
RL   Submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF157606; AAK68867.1; -.
DR   InterPro; IPR005594; YadA.
DR   Pfam; PF03895; YadA; 1.
SQ   SEQUENCE   591 AA;  62048 MW;  C0DC600798859C65 CRC64;

  Query Match             98.7%;  Score 2720;  DB 2;  Length 591;
  Best Local Similarity   99.1%;  Pred. No. 2.3e-107;
  Matches  535;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 NNEEQEEYLYLHPVQRTVAVLIVNSDKEGAGEKEKVEENSDWAVYFNEKGVLTAREITLK 60
              ||||||| ||| || |||||||||||||| ||||||||||||||||||||||||||||||
Db         52 NNEEQEEDLYLDPVLRTVAVLIVNSDKEGTGEKEKVEENSDWAVYFNEKGVLTAREITLK 111

Qy         61 AGDNLKIKQNGTNFTYSLKKDLTDLTSVGTEKLSFSAHGNKVNITSDTKGLNFAKETAGT 120
              ||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||
Db        112 AGDNLKIKQNGTNFTYSLKKDLTDLTSVGTEKLSFSANGNKVNITSDTKGLNFAKETAGT 171

Qy        121 NGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        172 NGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPG 231

Qy        181 TTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        232 TTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTG 291

Qy        241 KDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        292 KDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFA 351

Qy        301 SGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSP 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        352 SGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSP 411

Qy        361 SKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGDALNVG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        412 SKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGDALNVG 471

Qy        421 SKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLV 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        472 SKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLV 531

Qy        481 QAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        532 QAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 591


RESULT 4
Q9AQF0
ID   Q9AQF0      PRELIMINARY;      PRT;   592 AA.
AC   Q9AQF0;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Outer membrane protein.
GN   NHHA.
OS   Neisseria meningitidis.
OC   Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC   Neisseriaceae; Neisseria.
OX   NCBI_TaxID=487;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MC58;
RA   Peak I.R., Srikhanta Y., Dieckelman M., Moxon E.R., Jennings M.P.;
RT   "Identification and characterisation of a gene encoding a novel outer
RT   membrane protein of Neisseria meningitidis.";
RL   Submitted (FEB-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF125375; AAK09243.1; -.
DR   InterPro; IPR005594; YadA.
DR   Pfam; PF03895; YadA; 1.
SQ   SEQUENCE   592 AA;  62290 MW;  168986A97381EFC5 CRC64;

  Query Match             97.8%;  Score 2695.5;  DB 2;  Length 592;
  Best Local Similarity   98.2%;  Pred. No. 2.5e-106;
  Matches  531;  Conservative    4;  Mismatches    5;  Indels    1;  Gaps    1;

Qy          1 NNEE-QEEYLYLHPVQRTVAVLIVNSDKEGAGEKEKVEENSDWAVYFNEKGVLTAREITL 59
              |||  ::: ||| ||||||||||||||| ||||||||||||||||||||||||||||||
Db         52 NNERPRKKDLYLDPVQRTVAVLIVNSDKEGTGEKEKVEENSDWAVYFNEKGVLTAREITL 111

Qy         60 KAGDNLKIKQNGTNFTYSLKKDLTDLTSVGTEKLSFSAHGNKVNITSDTKGLNFAKETAG 119
              ||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||
Db        112 KAGDNLKIKQNGTNFTYSLKKDLTDLTSVGTEKLSFSANGNKVNITSDTKGLNFAKETAG 171

Qy        120 TNGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKP 179
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        172 TNGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKP 231

Qy        180 GTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVT 239
              |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        232 GTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGVKTSVIKEKDGKLVT 291

Qy        240 GKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTF 299
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        292 GKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTF 351

Qy        300 ASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVS 359
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        352 ASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVS 411

Qy        360 PSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGDALNV 419
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        412 PSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGDALNV 471

Qy        420 GSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGL 479
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        472 GSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGL 531

Qy        480 VQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQ 539
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        532 VQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQ 591

Qy        540 W 540
              |
Db        592 W 592


RESULT 5
Q9JPS3
ID   Q9JPS3      PRELIMINARY;      PRT;   590 AA.
AC   Q9JPS3;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Outer membrane protein GNA992.
GN   GNA992.
OS   Neisseria meningitidis.
OC   Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC   Neisseriaceae; Neisseria.
OX   NCBI_TaxID=487;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NGE28;
RX   MEDLINE=20175756; PubMed=10710308;
RA   Pizza M., Scarlato V., Masignani V., Giuliani M.M., Arico' B.,
```

```
RA   Comanducci M., Jennings G.T., Baldi L., Bartolini E., Capecchi B.,
RA   Galeotti C.L., Luzzi E., Manetti R., Marchetti E., Mora M., Nuti S.,
RA   Ratti G., Santini L., Savino S., Scarselli M., Storni E., Zuo P.,
RA   Broeker M., Hundt E., Knapp B., Blair E., Mason T., Tettelin H.,
RA   Hood D.W., Jeffries A.C., Saunders N.J., Granoff D.M., Venter C.,
RA   Moxon E.R., Grandi G., Rappuoli R.;
RT   "Identification of Vaccine Candidates Against Serogroup B
RT   Meningococcus by Whole-Genome Sequencing.";
RL   Science 287:1816-1820(2000).
DR   EMBL; AF226378; AAF42527.1; -.
DR   InterPro; IPR005594; YadA.
DR   Pfam; PF03895; YadA; 1.
SQ   SEQUENCE   590 AA;  61661 MW;  8AA476AC300D80C8 CRC64;

  Query Match           95.2%;  Score 2624.5;  DB 2;  Length 590;
  Best Local Similarity   96.7%;  Pred. No. 2.4e-103;
  Matches  522;  Conservative    4;  Mismatches   11;  Indels    3;  Gaps    2;

Qy          2 NEEQEEEYLYLHPVQRTVAVLIVNSDKEGAGEKEKVEENSDWAVYFNEKGVLTAREITLKA 61
              :|| ||   | ||||||||||||||| |||||||||||||||||||||||  |||||
Db         53 DEEDEE--DLDPVQRTVAVLIVNSDKEGTGEKEKVEENSDWAVYFNEKGVLTAGTITLKA 110

Qy         62 GDNLKIKQNGTNFTYSLKKDLTDLTSVGTEKLSFSAHGNKVNITSDTKGLNFAKETAGTN 121
              ||||||||||||||||||||||||||||||||||:|||||||||||||||||||||||
Db        111 GDNLKIKQNGTNFTYSLKKDLTDLTSVGTEKLSFSANGNKVNITSDTKGLNFAKETAGTN 170

Qy        122 GDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPGT 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        171 GDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPGT 230

Qy        182 TASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGK 241
              ||||||||||||||||||||||||||||||||||:|||||||||||||||||||||||
Db        231 TASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKRTEVKIGAKTSVIKEKDGKLVTGK 290

Qy        242 DKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFAS 301
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        291 GKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFAS 350

Qy        302 GKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPS 361
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        351 GKGTTATVSKDDQGNITVKYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPS 410

Qy        362 KGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGD-ALNVG 420
              |||||||||||||||||||||||||||||||||||||||||||||||||| : |||||
Db        411 KGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGALNVG 470

Qy        421 SKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLV 480
              ||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        471 SKDANKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLV 530

Qy        481 QAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        531 QAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 590


RESULT 6
Q93QY4
ID   Q93QY4      PRELIMINARY;      PRT;   594 AA.
AC   Q93QY4;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   NhhA outer membrane protein.
GN   NHHA.
OS   Neisseria meningitidis.
OC   Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC   Neisseriaceae; Neisseria.
OX   NCBI_TaxID=487;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=EG327;
RA   Peak I.R., Srikhanta Y., Dieckelman M., Moxon R., Jennings M.P.;
RT   "Identification and characterization of a gene encoding a novel outer
RT   membrane protein of Neisseria meningitidis.";
RL   Submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF157605; AAK68866.1; -.
DR   InterPro; IPR005594; YadA.
DR   Pfam; PF03895; YadA; 1.
SQ   SEQUENCE   594 AA;  62297 MW;  9DDD48B04B3A8EA2 CRC64;

  Query Match           91.1%;  Score 2512;  DB 2;  Length 594;
  Best Local Similarity   92.4%;  Pred. No. 1.3e-98;
  Matches  501;  Conservative   13;  Mismatches   20;  Indels    8;  Gaps    3;

Qy          6 EEYLYLHPVQRTVAVLIVNSDKEGAGEKEKVEENSDWAVYFNEKGVLTAREITLKAGDNL 65
              :: ||| ||||| ||   ||||| |||| | |:|:| |||::|||||| ||||||||
Db         54 DDDLYLEPVQRTAVVLSFRSDKEGTGEKE-VTEDSNWGVYFDKKGVLTAGTITLKAGDNL 112

Qy         66 KIKQ------NGTNFTYSLKKDLTDLTSVGTEKLSFSAHGNKVNITSDTKGLNFAKETAG 119
              ||||      | ::||||||||||||||||||||||: ||||||||||||||:||
Db        113 KIKQNTNENTNASSFTYSLKKDLTDLTSVGTEKLSFSANSNKVNITSDTKGLNFAKKTAE 172

Qy        120 TNGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKP 179
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        173 TNGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKP 232

Qy        180 GTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVT 239
              ||||||||||||||||||||||||||||||||||:|||||||||||||||||||||||
Db        233 GTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKRTEVKIGAKTSVIKEKDGKLVT 292

Qy        240 GKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTF 299
              ||||||| ||||:|||||||||||||||||||||||||||||||||||||||||||||||
Db        293 GKDKGENDSSTDKGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTF 352

Qy        300 ASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVS 359
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        353 ASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVS 412

Qy        360 PSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGD-ALN 418
              |||||||||||||||||||||||||||||||||||||||||||||||||||| : |||
Db        413 PSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGALN 472

Qy        419 VGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAG 478
              ||||  |||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        473 VGSKDANKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNHIDNVDGNARAGIAQAIATAG 532

Qy        479 LVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGY 538
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        533 LVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGY 592

Qy        539 QW 540
              ||
Db        593 QW 594


RESULT 7
Q9JPS6
ID   Q9JPS6      PRELIMINARY;      PRT;   600 AA.
AC   Q9JPS6;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Outer membrane protein GNA992.
GN   GNA992.
OS   Neisseria meningitidis.
OC   Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC   Neisseriaceae; Neisseria.
OX   NCBI_TaxID=487;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=E26;
RX   MEDLINE=20175756; PubMed=10710308;
RA   Pizza M., Scarlato V., Masignani V., Giuliani M.M., Arico' B.,
RA   Comanducci M., Jennings G.T., Baldi L., Bartolini E., Capecchi B.,
RA   Galeotti C.L., Luzzi E., Manetti R., Marchetti E., Mora M., Nuti S.,
RA   Ratti G., Santini L., Savino S., Scarselli M., Storni E., Zuo P.,
```

```
RA   Broeker M., Hundt E., Knapp B., Blair E., Mason T., Tettelin H.,
RA   Hood D.W., Jeffries A.C., Saunders N.J., Granoff D.M., Venter C.,
RA   Moxon E.R., Grandi G., Rappuoli R.;
RT   "Identification of Vaccine Candidates Against Serogroup B
RT   Meningococcus by Whole-Genome Sequencing.";
RL   Science 287:1816-1820(2000).
DR   EMBL; AF226371; AAF42520.1; -.
DR   InterPro; IPR005594; YadA.
DR   Pfam; PF03895; YadA; 1.
SQ   SEQUENCE   600 AA;  62762 MW;  36256963E0598CD1 CRC64;

  Query Match             91.1%;  Score 2511;  DB 2;  Length 600;
  Best Local Similarity   91.6%;  Pred. No. 1.5e-98;
  Matches  500;  Conservative   12;  Mismatches   24;  Indels   10;  Gaps    3;

Qy          1 NNEEQEEYLYLHPVQRTVAVLIVNSDKEGAGEKEKVEENSDWAVYFNEKGVLTAREITLK 60
              :|||:|   || || ||  ||   || |  |||| | ||::| :||:: ||: |  ||||
Db         59 DNEEEE---YLEPVVRTAPVLSFYSDAEDTGEKE-VTENTNWGIYFDKNGVIKAGTITLK 114

Qy         61 AGDNLKIKQ------NGTNFTYSLKKDLTDLTSVGTEKLSFSAHGNKVNITSDTKGLNFA 114
              |||||||||      | ::|||||||:|||||||||||||| |:|||||||||||||||
Db        115 AGDNLKIKQNTDENTNASSFTYSLKKELTDLTSVGTEKLSFGANGNKVNITSDTKGLNFA 174

Qy        115 KETAGTNGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNI 174
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        175 KETAGTNGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNI 234

Qy        175 KGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKD 234
              ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        235 KGVKPGTTASDNVDFVHTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKD 294

Qy        235 GKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSG 294
              ||||||| |||||||||||||||||||||||||||||||||||| ||||||||||||
Db        295 GKLVTGKGKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQKGQADKFETVTSG 354

Qy        295 TNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVI 354
              |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db        355 TNVTFASGKGTTATVSKDDQGNITVKYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVI 414

Qy        355 SGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDG 414
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        415 SGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDG 474

Qy        415 DALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAI 474
              ||||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        475 DALNVGSKDANKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAI 534

Qy        475 ATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASA 534
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        535 ATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASA 594

Qy        535 SVGYQW 540
              ||||||
Db        595 SVGYQW 600


RESULT 8
Q9JPI3
ID   Q9JPI3      PRELIMINARY;      PRT;   594 AA.
AC   Q9JPI3;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Outer membrane protein GNA992.
GN   GNA992.
OS   Neisseria meningitidis.
OC   Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC   Neisseriaceae; Neisseria.
OX   NCBI_TaxID=487;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NG3/88, and BZ232;
RX   MEDLINE=20175756; PubMed=10710308;
RA   Pizza M., Scarlato V., Masignani V., Giuliani M.M., Arico' B.,
RA   Comanducci M., Jennings G.T., Baldi L., Bartolini E., Capecchi B.,
RA   Galeotti C.L., Luzzi E., Manetti R., Marchetti E., Mora M., Nuti S.,
RA   Ratti G., Santini L., Savino S., Scarselli M., Storni E., Zuo P.,
RA   Broeker M., Hundt E., Knapp B., Blair E., Mason T., Tettelin H.,
RA   Hood D.W., Jeffries A.C., Saunders N.J., Granoff D.M., Venter C.,
RA   Moxon E.R., Grandi G., Rappuoli R.;
RT   "Identification of Vaccine Candidates Against Serogroup B
RT   Meningococcus by Whole-Genome Sequencing.";
RL   Science 287:1816-1820(2000).
DR   EMBL; AF226376; AAF42525.1; -.
DR   EMBL; AF226369; AAF42518.1; -.
DR   InterPro; IPR005594; YadA.
DR   Pfam; PF03895; YadA; 1.
SQ   SEQUENCE   594 AA;  62086 MW;  1B25E03B90D04B46 CRC64;

  Query Match             90.7%;  Score 2501;  DB 2;  Length 594;
  Best Local Similarity   92.1%;  Pred. No. 3.9e-98;
  Matches  499;  Conservative   12;  Mismatches   23;  Indels    8;  Gaps    3;

Qy          6 EEYLYLHPVQRTVAVLIVNSDKEGAGEKEKVEENSDWAVYFNEKGVLTAREITLKAGDNL 65
              :: ||| |||||  ||  ::| || |||| | |:|:| |||::|||||  |||||||||
Db         54 DDDLYLEPVQRTAPVLSFHADSEGTGEKE-VTEDSNWGVYFDKKGVLTAGTITLKAGDNL 112

Qy         66 KIKQ------NGTNFTYSLKKDLTDLTSVGTEKLSFSAHGNKVNITSDTKGLNFAKETAG 119
              ||||      | ::|||||||||||||| |||||| |:| ||||||||||||||||||
Db        113 KIKQNTDENTNASSFTYSLKKDLTDLTSVETEKLSFGANGKKVNITSDTKGLNFAKETAG 172

Qy        120 TNGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKP 179
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        173 TNGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKP 232

Qy        180 GTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVT 239
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        233 GTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVT 292

Qy        240 GKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTF 299
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db        293 GKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTKVTF 352

Qy        300 ASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVS 359
              ||| |||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        353 ASGNGTTATVSKDDQGNITVKYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVS 412

Qy        360 PSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGD-ALN 418
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||| : |||
Db        413 PSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGALN 472

Qy        419 VGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAG 478
              ||||  |||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        473 VGSKDANKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNHIDNVDGNARAGIAQAIATAG 532

Qy        479 LVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGY 538
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        533 LVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGY 592

Qy        539 QW 540
              ||
Db        593 QW 594


RESULT 9
Q9JPS2
ID   Q9JPS2      PRELIMINARY;      PRT;   594 AA.
AC   Q9JPS2;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Outer membrane protein GNA992.
GN   GNA992.
OS   Neisseria meningitidis.
OC   Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC   Neisseriaceae; Neisseria.
```

```
OX   NCBI_TaxID=487;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NGE31;
RX   MEDLINE=20175756; PubMed=10710308;
RA   Pizza M., Scarlato V., Masignani V., Giuliani M.M., Arico' B.,
RA   Comanducci M., Jennings G.T., Baldi L., Bartolini E., Capecchi B.,
RA   Galeotti C.L., Luzzi E., Manetti R., Marchetti E., Mora M., Nuti S.,
RA   Ratti G., Santini L., Savino S., Scarselli M., Storni E., Zuo P.,
RA   Broeker M., Hundt E., Knapp B., Blair E., Mason T., Tettelin H.,
RA   Hood D.W., Jeffries A.C., Saunders N.J., Granoff D.M., Venter C.,
RA   Moxon E.R., Grandi G., Rappuoli R.;
RT   "Identification of Vaccine Candidates Against Serogroup B
RT   Meningococcus by Whole-Genome Sequencing.";
RL   Science 287:1816-1820(2000).
DR   EMBL; AF226379; AAF42528.1; -.
DR   InterPro; IPR005594; YadA.
DR   Pfam; PF03895; YadA; 1.
SQ   SEQUENCE   594 AA;  62114 MW;  1E2A63A78F53D256 CRC64;

  Query Match             90.7%;  Score 2501;  DB 2;  Length 594;
  Best Local Similarity   92.1%;  Pred. No. 3.9e-98;
  Matches  499;  Conservative   12;  Mismatches   23;  Indels    8;  Gaps    3;

Qy          6 EEYLYLHPVQRTVAVLIVNSDKEGAGEKEKVEENSDWAVYFNEKGVLTAREITLKAGDNL 65
              :: ||| |||||  ||  ::| || |||| | |:|:| |||::|||||  |||||||||
Db         54 DDDLYLEPVQRTAPVLSFHADSEGTGEKE-VTEDSNWGVYFDKKGVLTAGTITLKAGDNL 112

Qy         66 KIKQ------NGTNFTYSLKKDLTDLTSVGTEKLSFSAHGNKVNITSDTKGLNFAKETAG 119
              ||||      | ::|||||||||||||||| |||||| |:| ||||||||||||||||||
Db        113 KIKQNTDENTNASSFTYSLKKDLTDLTSVETEKLSFGANGKKVNITSDTKGLNFAKETAG 172

Qy        120 TNGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKP 179
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        173 TNGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKP 232

Qy        180 GTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVT 239
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        233 GTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVT 292

Qy        240 GKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTF 299
              ||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db        293 GKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTKVTF 352

Qy        300 ASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVS 359
              ||| |||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        353 ASGNGTTATVSKDDQGNITVKYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVS 412

Qy        360 PSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGD-ALN 418
              |||||||||||||||||||||||||||||||||||||||||||||||||||||| : |||
Db        413 PSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGALN 472

Qy        419 VGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAG 478
              ||||  |||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        473 VGSKDANKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNHIDNVDGNARAGIAQAIATAG 532

Qy        479 LVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGY 538
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        533 LVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGY 592

Qy        539 QW 540
              ||
Db        593 QW 594


RESULT 10
Q9JPH7
ID   Q9JPH7      PRELIMINARY;      PRT;   594 AA.
AC   Q9JPH7;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Outer membrane protein GNA992 (NhhA outer membrane protein).
GN   GNA992 OR NHHA.
OS   Neisseria meningitidis.
OC   Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC   Neisseriaceae; Neisseria.
OX   NCBI_TaxID=487;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BZ198, and 297-0;
RX   MEDLINE=20175756; PubMed=10710308;
RA   Pizza M., Scarlato V., Masignani V., Giuliani M.M., Arico' B.,
RA   Comanducci M., Jennings G.T., Baldi L., Bartolini E., Capecchi B.,
RA   Galeotti C.L., Luzzi E., Manetti R., Marchetti E., Mora M., Nuti S.,
RA   Ratti G., Santini L., Savino S., Scarselli M., Storni E., Zuo P.,
RA   Broeker M., Hundt E., Knapp B., Blair E., Mason T., Tettelin H.,
RA   Hood D.W., Jeffries A.C., Saunders N.J., Granoff D.M., Venter C.,
RA   Moxon E.R., Grandi G., Rappuoli R.;
RT   "Identification of Vaccine Candidates Against Serogroup B
RT   Meningococcus by Whole-Genome Sequencing.";
RL   Science 287:1816-1820(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BZ198;
RA   Peak I.R., Srikhanta Y., Dieckelman M., Moxon R., Jennings M.P.;
RT   "Identification and characterization of a gene encoding a novel outer
RT   membrane protein of Neisseria meningitidis.";
RL   Submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF226368; AAF42517.1; -.
DR   EMBL; AF226358; AAF42507.1; -.
DR   EMBL; AF157604; AAK68865.1; -.
DR   InterPro; IPR005594; YadA.
DR   Pfam; PF03895; YadA; 1.
SQ   SEQUENCE   594 AA;  62361 MW;  436BDDED68263C5C CRC64;

  Query Match             90.4%;  Score 2493;  DB 2;  Length 594;
  Best Local Similarity   91.6%;  Pred. No. 8.5e-98;
  Matches  500;  Conservative    9;  Mismatches   29;  Indels    8;  Gaps    3;

Qy          2 NEEQEEYLYLHPVQRTVAVLIVNSDKEGAGEKEKVEENSDWAVYFNEKGVLTAREITLKA 61
              |   :: ||| |||||  ||   ||||| |||| |: |:|||||:|| || |  |||||
Db         50 NATDDDDLYLEPVQRTAVVLSFRSDKEGTGEKEGTED-SNWAVYFDEKRVLKAGAITLKA 108

Qy         62 GDNLKIKQ------NGTNFTYSLKKDLTDLTSVGTEKLSFSAHGNKVNITSDTKGLNFAK 115
              ||||||||      | ::|||||||||||||||| |||||| |:|||||||||||||||
Db        109 GDNLKIKQNTNENTNDSSFTYSLKKDLTDLTSVETEKLSFGANGNKVNITSDTKGLNFAK 168

Qy        116 ETAGTNGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIK 175
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        169 ETAGTNGDPTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIK 228

Qy        176 GVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDG 235
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        229 GVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDG 288

Qy        236 KLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGT 295
              |||||| | ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        289 KLVTGKGKDENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGT 348

Qy        296 NVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVIS 355
              ||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        349 NVTFASGKGTTATVSKDDQGNITVKYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVIS 408

Qy        356 GNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGD 415
              |||||||||||||||||||||||||||||||| |||||||||||||||||||| :
Db        409 GNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMAPQFSSVSLGAGADAPTLSVDDE 468

Qy        416 -ALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAI 474
               |||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        469 GALNVGSKDTNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAI 528

Qy        475 ATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASA 534
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        529 ATAGLVQAYLPGKSMMAIGGDTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASA 588
```

```
Qy        535 SVGYQW 540
              ||||||
Db        589 SVGYQW 594


RESULT 11
Q9JPR9
ID   Q9JPR9      PRELIMINARY;      PRT;   598 AA.
AC   Q9JPR9;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Outer membrane protein GNA992.
GN   GNA992.
OS   Neisseria meningitidis.
OC   Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC   Neisseriaceae; Neisseria.
OX   NCBI_TaxID=487;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NGH36;
RX   MEDLINE=20175756; PubMed=10710308;
RA   Pizza M., Scarlato V., Masignani V., Giuliani M.M., Arico' B.,
RA   Comanducci M., Jennings G.T., Baldi L., Bartolini E., Capecchi B.,
RA   Galeotti C.L., Luzzi E., Manetti R., Marchetti E., Mora M., Nuti S.,
RA   Ratti G., Santini L., Savino S., Scarselli M., Storni E., Zuo P.,
RA   Broeker M., Hundt E., Knapp B., Blair E., Mason T., Tettelin H.,
RA   Hood D.W., Jeffries A.C., Saunders N.J., Granoff D.M., Venter C.,
RA   Moxon E.R., Grandi G., Rappuoli R.;
RT   "Identification of Vaccine Candidates Against Serogroup B
RT   Meningococcus by Whole-Genome Sequencing.";
RL   Science 287:1816-1820(2000).
DR   EMBL; AF226382; AAF42531.1; -.
DR   InterPro; IPR005594; YadA.
DR   Pfam; PF03895; YadA; 1.
SQ   SEQUENCE   598 AA;  62718 MW;  9095F8E31AD7C76D CRC64;

  Query Match            90.4%;  Score 2491;  DB 2;  Length 598;
  Best Local Similarity  90.9%;  Pred. No. 1e-97;
  Matches  500;  Conservative    9;  Mismatches   29;  Indels   12;  Gaps    3;

Qy          2 NEEQEEYLYLHPVQRTVAVLIVNSDKEGAGEKEKVEENSDWAVYFNEKGVLTAREITLKA 61
              |   :: ||| ||||| ||   ||||| |||| |: |:|||||:|| || |  |||||
Db         50 NATDDDDLYLEPVQRTAVVLSFRSDKEGTGEKEGTED-SNWAVYFDEKRVLKAGAITLKA 108

Qy         62 GDNLKIKQ----------NGTNFTYSLKKDLTDLTSVGTEKLSFSAHGNKVNITSDTKGL 111
              ||||||||          | ::||||||||||||||| |||||| |:||||||||||||
Db        109 GDNLKIKQNTNENTNENTNDSSFTYSLKKDLTDLTSVETEKLSFGANGNKVNITSDTKGL 168

Qy        112 NFAKETAGTNGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAG 171
              |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        169 NFAKETAGTNGDPTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAG 228

Qy        172 WNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIK 231
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        229 WNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIK 288

Qy        232 EKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETV 291
              |||||||||| | ||||||||||||||||||||||||||||||||||||||||||||||
Db        289 EKDGKLVTGKGKDENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETV 348

Qy        292 TSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSG 351
              |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db        349 TSGTNVTFASGKGTTATVSKDDQGNITVKYDVNVGDALNVNQLQNSGWNLDSKAVAGSSG 408

Qy        352 KVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLS 411
              |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        409 KVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMAPQFSSVSLGAGADAPTLS 468

Qy        412 VDGD-ALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGI 470
              || : |||||||  |||||||||||||||||||||||||||||||||||||||||||| |||
Db        469 VDDEGALNVGSKDTNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNALAGI 528

Qy        471 AQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHF 530
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        529 AQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHF 588

Qy        531 GASASVGYQW 540
              ||||||||||
Db        589 GASASVGYQW 598


RESULT 12
Q9JPS0
ID   Q9JPS0      PRELIMINARY;      PRT;   598 AA.
AC   Q9JPS0;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Outer membrane protein GNA992 (NhhA outer membrane protein).
GN   GNA992 OR NHHA.
OS   Neisseria meningitidis.
OC   Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC   Neisseriaceae; Neisseria.
OX   NCBI_TaxID=487;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NGH15;
RX   MEDLINE=20175756; PubMed=10710308;
RA   Pizza M., Scarlato V., Masignani V., Giuliani M.M., Arico' B.,
RA   Comanducci M., Jennings G.T., Baldi L., Bartolini E., Capecchi B.,
RA   Galeotti C.L., Luzzi E., Manetti R., Marchetti E., Mora M., Nuti S.,
RA   Ratti G., Santini L., Savino S., Scarselli M., Storni E., Zuo P.,
RA   Broeker M., Hundt E., Knapp B., Blair E., Mason T., Tettelin H.,
RA   Hood D.W., Jeffries A.C., Saunders N.J., Granoff D.M., Venter C.,
RA   Moxon E.R., Grandi G., Rappuoli R.;
RT   "Identification of Vaccine Candidates Against Serogroup B
RT   Meningococcus by Whole-Genome Sequencing.";
RL   Science 287:1816-1820(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=H15;
RA   Peak I.R., Srikhanta Y., Dieckelman M., Moxon R., Jennings M.P.;
RT   "Identification and characterization of a gene encoding a novel outer
RT   membrane protein of Neisseria meningitidis.";
RL   Submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF226381; AAF42530.1; -.
DR   EMBL; AF157607; AAK68868.1; -.
DR   InterPro; IPR005594; YadA.
DR   Pfam; PF03895; YadA; 1.
SQ   SEQUENCE   598 AA;  62763 MW;  E6C7AEF0BB8A63CB CRC64;

  Query Match            89.9%;  Score 2478;  DB 2;  Length 598;
  Best Local Similarity  90.4%;  Pred. No. 3.7e-97;
  Matches  497;  Conservative   10;  Mismatches   31;  Indels   12;  Gaps    3;

Qy          2 NEEQEEYLYLHPVQRTVAVLIVNSDKEGAGEKEKVEENSDWAVYFNEKGVLTAREITLKA 61
              |   :: ||| ||||| ||   ||||| |||| |: |:|||||:|| || |  |||||
Db         50 NATDDDDLYLEPVQRTAVVLSFRSDKEGTGEKEGTED-SNWAVYFDEKRVLKAGAITLKA 108

Qy         62 GDNLKIKQ----------NGTNFTYSLKKDLTDLTSVGTEKLSFSAHGNKVNITSDTKGL 111
              ||||||||          | ::||||||||||||||| |||||| |:||||||||||||
Db        109 GDNLKIKQNTNENTNENTNDSSFTYSLKKDLTDLTSVETEKLSFGANGNKVNITSDTKGL 168

Qy        112 NFAKETAGTNGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAG 171
              |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        169 NFAKETAGTNGDPTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAG 228

Qy        172 WNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIK 231
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        229 WNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIK 288

Qy        232 EKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETV 291
              |||||||||| | ||||||||||||||||||||||||||||||||||||||||||||||
Db        289 EKDGKLVTGKGKDENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETV 348
```

```
Qy        292 TSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSG 351
              |||| |||||| |||||||||||||||| |||||||||||||||||||||||||||||||
Db        349 TSGTKVTFASGNGTTATVSKDDQGNITVKYDVNVGDALNVNQLQNSGWNLDSKAVAGSSG 408

Qy        352 KVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLS 411
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        409 KVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLS 468

Qy        412 VDGD-ALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGI 470
              || : |||||||  |||||||||||||||||||||||||||||||||||||||||||||||
Db        469 VDDEGALNVGSKDANKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGI 528

Qy        471 AQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHF 530
              ||||||||| |||||||||||||||||||||||||||||||||| |||:||||||||||||
Db        529 AQAIATAGLAQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDTGNWVIKGTASGNSRGHF 588

Qy        531 GASASVGYQW 540
              ||||||||||
Db        589 GASASVGYQW 598


RESULT 13
Q93QY5
ID   Q93QY5      PRELIMINARY;      PRT;   598 AA.
AC   Q93QY5;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   NhhA outer membrane protein.
GN   NHHA.
OS   Neisseria meningitidis.
OC   Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC   Neisseriaceae; Neisseria.
OX   NCBI_TaxID=487;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BZ10;
RA   Peak I.R., Srikhanta Y., Dieckelman M., Moxon R., Jennings M.P.;
RT   "Identification and characterization of a gene encoding a novel outer
RT   membrane protein of Neisseria meningitidis.";
RL   Submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF157603; AAK68864.1; -.
DR   InterPro; IPR005594; YadA.
DR   Pfam; PF03895; YadA; 1.
SQ   SEQUENCE   598 AA;  62687 MW;  18CEFFE6410A15DF CRC64;

  Query Match             89.9%;  Score 2478;  DB 2;  Length 598;
  Best Local Similarity   90.2%;  Pred. No. 3.7e-97;
  Matches  496;  Conservative   11;  Mismatches   31;  Indels   12;  Gaps    3;

Qy          2 NEEQEEYLYLHPVQRTVAVLIVNSDKEGAGEKEKVEENSDWAVYFNEKGVLTAREITLKA 61
              |   :: ||| |||||  ||   ||||| ||||  |: |:|||||:|| || |  |||||
Db         50 NATDDDDLYLEPVQRTAVVLSFRSDKEGTGEKEGTED-SNWAVYFDEKRVLKAGAITLKA 108

Qy         62 GDNLKIKQ----------NGTNFTYSLKKDLTDLTSVGTEKLSFSAHGNKVNITSDTKGL 111
              ||||||||          | ::||||||||||||||| |||||| |:||||||||||||
Db        109 GDNLKIKQNTNENTNENTNDSSFTYSLKKDLTDLTSVETEKLSFGANGNKVNITSDTKGL 168

Qy        112 NFAKETAGTNGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAG 171
              |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db        169 NFAKETAGTNGDPTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAG 228

Qy        172 WNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIK 231
              ||||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||
Db        229 WNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKRTEVKIGAKTSVIK 288

Qy        232 EKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETV 291
              |||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        289 EKDGKLVTGKGKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETV 348

Qy        292 TSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSG 351
              |||| |||||| |||||||||||||||| |||||||||||||||||||||||||||||||
Db        349 TSGTKVTFASGNGTTATVSKDDQGNITVKYDVNVGDALNVNQLQNSGWNLDSKAVAGSSG 408

Qy        352 KVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLS 411
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        409 KVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLS 468

Qy        412 VDGD-ALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGI 470
              || : |||||||  |||||||||||||||||||||||||||||||||||||||||||||||
Db        469 VDDEGALNVGSKDANKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGI 528

Qy        471 AQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHF 530
              ||||||||| |||||||||||||||||||||||||||||||||| |||:||||||||||||
Db        529 AQAIATAGLAQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDTGNWVIKGTASGNSRGHF 588

Qy        531 GASASVGYQW 540
              ||||||||||
Db        589 GTSASVGYQW 598


RESULT 14
Q9JPT0
ID   Q9JPT0      PRELIMINARY;      PRT;   598 AA.
AC   Q9JPT0;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Outer membrane protein GNA992.
GN   GNA992.
OS   Neisseria meningitidis.
OC   Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC   Neisseriaceae; Neisseria.
OX   NCBI_TaxID=487;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=2996;
RX   MEDLINE=20175756; PubMed=10710308;
RA   Pizza M., Scarlato V., Masignani V., Giuliani M.M., Arico' B.,
RA   Comanducci M., Jennings G.T., Baldi L., Bartolini E., Capecchi B.,
RA   Galeotti C.L., Luzzi E., Manetti R., Marchetti E., Mora M., Nuti S.,
RA   Ratti G., Santini L., Savino S., Scarselli M., Storni E., Zuo P.,
RA   Broeker M., Hundt E., Knapp B., Blair E., Mason T., Tettelin H.,
RA   Hood D.W., Jeffries A.C., Saunders N.J., Granoff D.M., Venter C.,
RA   Moxon E.R., Grandi G., Rappuoli R.;
RT   "Identification of Vaccine Candidates Against Serogroup B
RT   Meningococcus by Whole-Genome Sequencing.";
RL   Science 287:1816-1820(2000).
DR   EMBL; AF226359; AAF42508.1; -.
DR   InterPro; IPR005594; YadA.
DR   Pfam; PF03895; YadA; 1.
SQ   SEQUENCE   598 AA;  62763 MW;  63A6A3BD7F0F2EE3 CRC64;

  Query Match             89.9%;  Score 2478;  DB 2;  Length 598;
  Best Local Similarity   90.2%;  Pred. No. 3.7e-97;
  Matches  496;  Conservative   11;  Mismatches   31;  Indels   12;  Gaps    3;

Qy          2 NEEQEEYLYLHPVQRTVAVLIVNSDKEGAGEKEKVEENSDWAVYFNEKGVLTAREITLKA 61
              |   :: ||| |||||  ||   ||||| ||||  |: |:|||||:|| || |  |||||
Db         50 NATDDDDLYLEPVQRTAVVLSFRSDKEGTGEKEGTED-SNWAVYFDEKRVLKAGAITLKA 108

Qy         62 GDNLKIKQ----------NGTNFTYSLKKDLTDLTSVGTEKLSFSAHGNKVNITSDTKGL 111
              ||||||||          | ::||||||||||||||| |||||| |:||||||||||||
Db        109 GDNLKIKQNTNENTNENTNDSSFTYSLKKDLTDLTSVETEKLSFGANGNKVNITSDTKGL 168

Qy        112 NFAKETAGTNGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAG 171
              |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db        169 NFAKETAGTNGDPTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAG 228

Qy        172 WNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIK 231
              ||||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||
Db        229 WNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKRTEVKIGAKTSVIK 288

Qy        232 EKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETV 291
              |||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        289 EKDGKLVTGKGKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETV 348
```

```
Qy        292 TSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSG 351
              |||| |||||| |||||||||||||||| |||||||||||||||||||||||||||||||
Db        349 TSGTKVTFASGNGTTATVSKDDQGNITVKYDVNVGDALNVNQLQNSGWNLDSKAVAGSSG 408

Qy        352 KVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLS 411
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        409 KVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLS 468

Qy        412 VDGD-ALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGI 470
              || : |||||||  ||||||||||||||||||||||||||||||||||||||||||||||||
Db        469 VDDEGALNVGSKDANKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGI 528

Qy        471 AQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHF 530
              ||||||||| |||||||||||||||||||||||||||||| |||:|||||||||||||||
Db        529 AQAIATAGLAQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDTGNWVIKGTASGNSRGHF 588

Qy        531 GASASVGYQW 540
              | ||||||||
Db        589 GTSASVGYQW 598


RESULT 15
Q9JPS4
ID   Q9JPS4      PRELIMINARY;      PRT;   526 AA.
AC   Q9JPS4;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Outer membrane protein GNA992.
GN   GNA992.
OS   Neisseria meningitidis.
OC   Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC   Neisseriaceae; Neisseria.
OX   NCBI_TaxID=487;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NG6/88;
RX   MEDLINE=20175756; PubMed=10710308;
RA   Pizza M., Scarlato V., Masignani V., Giuliani M.M., Arico' B.,
RA   Comanducci M., Jennings G.T., Baldi L., Bartolini E., Capecchi B.,
RA   Galeotti C.L., Luzzi E., Manetti R., Marchetti E., Mora M., Nuti S.,
RA   Ratti G., Santini L., Savino S., Scarselli M., Storni E., Zuo P.,
RA   Broeker M., Hundt E., Knapp B., Blair E., Mason T., Tettelin H.,
RA   Hood D.W., Jeffries A.C., Saunders N.J., Granoff D.M., Venter C.,
RA   Moxon E.R., Grandi G., Rappuoli R.;
RT   "Identification of Vaccine Candidates Against Serogroup B
RT   Meningococcus by Whole-Genome Sequencing.";
RL   Science 287:1816-1820(2000).
DR   EMBL; AF226377; AAF42526.1; -.
DR   InterPro; IPR005594; YadA.
DR   Pfam; PF03895; YadA; 1.
SQ   SEQUENCE   526 AA;  54732 MW;  F543C4234AF807CA CRC64;

  Query Match           88.6%;  Score 2442.5;  DB 2;  Length 526;
  Best Local Similarity  92.0%;  Pred. No. 9.9e-96;
  Matches  483;  Conservative   10;  Mismatches   21;  Indels   11;  Gaps    2;

Qy         26 DKEGAGEKEKVEENSDWAVYFNEKGVLTAREITLKAGDNLKIKQ----------NGTNFT 75
              |||| || |    |  |::|::    |    :|||||||||||          | ::||
Db          3 DKEGNGENEST-GNIGWSIYYDNHNTLHGATVTLKAGDNLKIKQNTNKNTNENTNDSSFT 61

Qy         76 YSLKKDLTDLTSVGTEKLSFSAHGNKVNITSDTKGLNFAKETAGTNGDTTVHLNGIGSTL 135
              |||||||||||||| |||||| |:||||||||||||||||||||||||||||||||||||
Db         62 YSLKKDLTDLTSVETEKLSFGANGNKVNITSDTKGLNFAKETAGTNGDTTVHLNGIGSTL 121

Qy        136 TDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDT 195
              ||||||||||||||||||:|||||||||||||||||||||||||||||| ||||
Db        122 TDTLLNTGATTNVTNDNVTDDKKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVHTYDT 181

Qy        196 VEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEG 255
              ||||||||||||||||||||:|||||||||||||||||||||| ||||||||||||
Db        182 VEFLSADTKTTTVNVESKDNGKRTEVKIGAKTSVIKEKDGKLVTGKGKGENGSSTDEGEG 241

Qy        256 LVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQG 315
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 LVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQG 301

Qy        316 NITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGN 375
              |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 NITVKYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGN 361

Qy        376 NIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGDALNVGSKKDNKPVRITNVAP 435
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 NIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGDALNVGSKKDNKPVRITNVAP 421

Qy        436 GVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGG 495
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        422 GVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGG 481

Qy        496 TYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 540
              |||||||||||||||||||||||||||||||||||||||||||||
Db        482 TYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 526


Search completed: October  6, 2003, 09:30:47
Job time : 41.661 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: October 6, 2003, 09:13:35 ; Search time 15.4863 Seconds
(without alignments)
3359.577 Million cell updates/sec

Title: US-09-771-382-34
Perfect score: 2735
Sequence: 1 TDEDEEEELESVQRSVVGSI.....TASGNSRGHFGASASVGYQW 541

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_76:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2527	92.4	592	2	A81888	probable surface f
2	2330.5	85.2	591	2	G81133	adhesin NMB0992 [i
3	435.5	15.9	298	2	I64138	adhesin homolog HI
4	392.5	14.4	2059	2	D82671	surface protein XF
5	377	13.8	1588	2	A86036	probable adhesin Z
6	377	13.8	1588	2	H91188	probable adhesin E
7	368.5	13.5	1190	2	A82615	surface protein XF
8	363	13.3	1107	2	AC0976	probable autotrans
9	321.5	11.8	658	2	AH0110	probable surface p
10	231	8.4	1536	2	A43855	high-molecular-wei
11	225	8.2	1109	2	A56143	surface-array prot
12	220.5	8.1	1004	2	C82672	surface-exposed ou
13	214	7.8	1910	2	AF0394	probable adhesin h
14	212.5	7.8	4919	2	T31105	hypothetical prote
15	210	7.7	2249	2	A41477	190K surface antig
16	207	7.6	3705	2	AD0123	probable autotrans
17	206.5	7.6	1655	2	E97835	hypothetical prote
18	204.5	7.5	5188	2	B85547	probable RTX famil
19	203.5	7.4	5291	2	F90696	hypothetical prote
20	200	7.3	1238	2	AH0038	probable exported
21	198	7.2	2020	2	C48399	ABC-type transport
22	197.5	7.2	936	2	I40711	sapB protein - Cam
23	197.5	7.2	4152	2	T31102	filamentous hemagg
24	197	7.2	1577	2	A35140	hemolysin A precur
25	197	7.2	3029	2	S76109	hypothetical prote
26	195	7.1	1477	2	B43855	high-molecular-wei
27	195	7.1	1651	2	JC1340	outer membrane pro
28	194	7.1	1275	2	T33369	hypothetical prote
29	193.5	7.1	2232	2	T34434	hypothetical prote
30	189.5	6.9	2021	2	A97859	190-KDa cell surfa
31	189	6.9	585	2	F85809	hypothetical prote
32	189	6.9	920	2	I40614	surface array prot
33	188.5	6.9	1268	2	B99789	hemagglutinin/hemo
34	188.5	6.9	1270	2	E85649	hypothetical prote
35	187.5	6.9	365	2	AB3486	cell surface prote
36	187	6.8	585	2	F90961	flagellin [importe
37	187	6.8	1428	2	AC2224	hypothetical prote
38	187	6.8	2554	2	AB3528	extracellular seri
39	186.5	6.8	1361	2	T03415	S-layer protein -
40	185.5	6.8	1430	2	AF0351	probable autotrans
41	185.5	6.8	1635	2	AI0452	hemolysin [importe
42	184.5	6.7	1975	2	B81192	hemagglutinin/hemo
43	184.5	6.7	1995	2	G81044	hemagglutinin/hemo
44	184	6.7	1385	2	D89824	hypothetical prote
45	184	6.7	2468	2	A83412	hypothetical prote

ALIGNMENTS

RESULT 1
A81888
probable surface fibril protein NMA1200 [imported] - Neisseria meningitidis (strain Z
C;Species: Neisseria meningitidis
C;Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001
C;Accession: A81888
R;Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Mo
; Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandre
Nature 404, 502-506, 2000
A;Title: Complete DNA sequence of a serogroup A strain of Neisseria menigitidis Z2491
A;Reference number: A81775; MUID:20222556; PMID:10761919
A;Accession: A81888
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-592 <PAR>
A;Cross-references: GB:AL162755; GB:AL157959; NID:g7379742; PIDN:CAB84461.1; PID:g737
A;Experimental source: serogroup A, strain Z2491
C;Genetics:
A;Gene: NMA1200

Query Match 92.4%; Score 2527; DB 2; Length 592;
Best Local Similarity 93.4%; Pred. No. 7e-121;
Matches 507; Conservative 9; Mismatches 23; Indels 4; Gaps 2;

```
Qy          1 TDEDEEEELESVQRSVVGSIQASMEGSVELETISLSMTNDSKEFVDPYIVVTLKAGDNLK 60
              ||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db         52 TDEDEEEELESVQRSVVGSIQASMEGSGELETISLSMTNDSKEFVDPYIVVTLKAGDNLK 111

Qy         61 IKQNTNENTNASSFTYSLKKDLTGLINVETEKLSFGANGKKVNIISDTKGLNFAKETAGT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        112 IKQNTNENTNASSFTYSLKKDLTGLINVETEKLSFGANGKKVNIISDTKGLNFAKETAGT 171

Qy        121 NGDTTVHLNGIGSTLTDMLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPG 180
              |||||||||||||||| |  : |:     :  |    ||||:||||||||||||| |
Db        172 NGDTTVHLNGIGSTLTDTLAGSSASHVDAGNQST--HYTRAASIKDVLNAGWNIKGVKTG 229

Qy        181 TTA--SDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLV 238
              :|   |:||||||||||||||||||||||||||||:||||||||||||||||||||||||
Db        230 STTGQSENVDFVRTYDTVEFLSADTKTTTVNVESKDNGKRTEVKIGAKTSVIKEKDGKLV 289

Qy        239 TGKGKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTKVT 298
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db        290 TGKGKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVT 349

Qy        299 FASGNGTTATVSKDDQGNITVKYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNV 358
              |||| ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        350 FASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNV 409

Qy        359 SPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGAL 418
              |||||||||||||||||||:|||||||||||| ||||||||||||||||||||||||||||
Db        410 SPSKGKMDETVNINAGNNIEISRNGKNIDIATSMAPQFSSVSLGAGADAPTLSVDDEGAL 469
```

```
Qy        419 NVGSKDANKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVNGNARAGIAQAIATA 478
              ||||||||||||||||||||||||||||||||||||||||||||||:||||||||||||||
Db        470 NVGSKDANKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATA 529

Qy        479 GLVQAYLPGKSMMAIGGGTYLGEAGYAIGYSSISAGGNWIIKGTASGNSRGHFGASASVG 538
              |||||||||||||||||| |||||||||||| ||||||||||||||||||||||||||||
Db        530 GLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVG 589

Qy        539 YQW 541
              |||
Db        590 YQW 592
```

RESULT 2
G81133
adhesin NMB0992 [imported] - Neisseria meningitidis (strain MC58 serogroup B)
C;Species: Neisseria meningitidis
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001
C;Accession: G81133
R;Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;
ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Masignani, V.; Pizza, M.
Science 287, 1809-1815, 2000
A;Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ve
A;Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A;Reference number: A81000; MUID:20175755; PMID:10710307
A;Accession: G81133
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-591 <TET>
A;Cross-references: GB:AE002450; GB:AE002098; NID:g7226229; PIDN:AAF41395.1; PID:g722623
A;Experimental source: serogroup B, strain MC58
C;Genetics:
A;Gene: NMB0992

  Query Match 85.2%; Score 2330.5; DB 2; Length 591;
  Best Local Similarity 86.8%; Pred. No. 6.3e-111;
  Matches 475; Conservative 22; Mismatches 35; Indels 15; Gaps 6;

```
Qy          2 DEDEEEE--LESVQRSV-VGSIQASMEGSVELETISLSMTNDSKEFVDPYIV----VTLK 54
              :|::||:  |: |||:| |  : :  ||: | | :    :: :  | :  ::    :|||
Db         53 NEEQEEDLYLDPVQRTVAVLIVNSDKEGTGEKEKVE-ENSDWAVYFNEKGVLTAREITLK 111

Qy         55 AGDNLKIKQNTNENTNASSFTYSLKKDLTGLINVETEKLSFGANGKKVNIISDTKGLNFA 114
              |||||||||    | ::||||||||||| | :| |||||| ||| |||| ||||||||||
Db        112 AGDNLKIKQ------NGTNFTYSLKKDLTDLTSVGTEKLSFSANGNKVNITSDTKGLNFA 165

Qy        115 KETAGTNGDTTVHLNGIGSTLTDMLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNI 174
              ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        166 KETAGTNGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNI 225

Qy        175 KGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKD 234
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        226 KGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKD 285

Qy        235 GKLVTGKGKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSG 294
              ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        286 GKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSG 345

Qy        295 TKVTFASGNGTTATVSKDDQGNITVKYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVI 354
              | |||||| ||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        346 TNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVI 405

Qy        355 SGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDD 414
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        406 SGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDG 465

Qy        415 EGALNVGSKDANKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVNGNARAGIAQA 474
              : |||||||  ||||||||||||||||||||||||||||||||||||:||||||||||
Db        466 D-ALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQA 524

Qy        475 IATAGLVQAYLPGKSMMAIGGGTYLGEAGYAIGYSSISAGGNWIIKGTASGNSRGHFGAS 534
              |||||||||||||||||||||| |||||||||||| ||||||||||||||||||||||||
Db        525 IATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGAS 584

Qy        535 ASVGYQW 541
              |||||||
Db        585 ASVGYQW 591
```

RESULT 3
I64138
adhesin homolog HI1732 - Haemophilus influenzae (strain Rd KW20)
C;Species: Haemophilus influenzae
C;Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 24-Oct-1997
C;Accession: I64138
R;Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage
; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman
, D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M.
Science 269, 496-512, 1995
A;Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Vente
A;Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
A;Reference number: A64000; MUID:95350630; PMID:7542800
A;Accession: I64138
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-298 <TIGR>
A;Cross-references: GB:U32846; GB:L42023; NID:g1574588; PID:g1574589; TIGR:HI1732

  Query Match 15.9%; Score 435.5; DB 2; Length 298;
  Best Local Similarity 48.0%; Pred. No. 2.6e-15;
  Matches 106; Conservative 27; Mismatches 57; Indels 31; Gaps 7;

```
Qy         51 VTLKAGDNLKIKQNTNENTNASSFTYSLKKDLTGLINVETEKLS----------FGANG 99
              :||||| ||| |     :    | |::| |||    :|:| |:|           |
Db         87 LTLKAGKNLKAKL----DQGGKSVTFALAKDL----DVKTAKVSDTLTIGGNTPAAGGAT 138

Qy        100 KKVNIISDTKGLNFAKETAGTNGDTTVHLNGIGSTLTDMLLNTGATTNVTNDNVTDDEKK 159
               ||:| |   ||  ||   |||||| ||||||: ||| |:  ||||:|:||  : :| ||
Db        139 PKVSITSTADGLKLAK---GTNGDTAVHLNGLASTLPDVTTNTGASTSVT-FSPSDIEKT 194

Qy        160 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKT 219
              |||::|||||||||||| |     ::||| |  || |||:: |  |  | : :|:||| |
Db        195 RAATIKDVLNAGWNIKGAKVAGGNTENVDLVAGYDNVEFITGDKNTLDVVLTAKENGKTT 254

Qy        220 EVKIGAKTSVIKEKDGKLVTGK-------GKGENGS-STDE 252
              |||   |||||: :|||:|||       |   | :  |||
Db        255 EVKFTPKTSVIKDNNGKLLTGKQLKDANTGTATNATEDTDE 295
```

RESULT 4
D82671
surface protein XF1529 [imported] - Xylella fastidiosa (strain 9a5c)
C;Species: Xylella fastidiosa
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C;Accession: D82671
R;anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Seq
Nature 406, 151-157, 2000
A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A;Reference number: A82515; MUID:20365717; PMID:10910347
A;Note: for a complete list of authors see reference number A59328 below
A;Accession: D82671
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-2059 <SIM>
A;Cross-references: GB:AE003982; GB:AE003849; NID:g9106554; PIDN:AAF84338.1; GSPDB:GN
A;Experimental source: strain 9a5c
R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer
as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Fr
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; La
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins
A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.
, F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri,

```
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silvei
M.; Tsuhako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A;Reference number: A59328
A;Contents: annotation
C;Genetics:
A;Gene: XF1529

  Query Match             14.4%;  Score 392.5;  DB 2;  Length 2059;
  Best Local Similarity   24.8%;  Pred. No. 4.7e-12;
  Matches  173;  Conservative   98;  Mismatches  225;  Indels  201;  Gaps   32;

Qy          7 EELESVQRSVVGSIQASMEGSVELETISLSMTNDSKEFVDPYIVVTLKAGD-NLKIKQNT 65
                || ::|  | : ||  :::   :  |:: |  | ||  |     | ||  | || | ::
Db       1402 EETDAVNFSQLKSISTAVD---QGWTLTASGANGSK--VASGGTVDLKNTDGNLTISKSG 1456

Qy         66 NENTNASSFTYSLKKDLTGLINVETEKLSFGANGKKV--NIISDTKGLNFAKETAGT--- 120
              : |     : :   ||:    : | : :|     :| ||  |:: |: |        :: :
Db       1457 DSNDVVFNLSEDLKEKSITVGNTQLDK-----DGVKVSSNVLLDSNELVITSHSSTSSVK 1511

Qy        121 ---NGDTTVH---LNGIGSTLTDML------------------LNTGA--TTNVT---- 149
                 ||:: |:   :|| |  : |::                   :| |:   ||||
Db       1512 TLANGESVVNRTVVNGDGVNIDDVVVVNDLGLSIVGGASLTLSGINAGSHKITNVTAGTE 1571

Qy        150 -NDNVTDDEKKRAASVKDVLNAGWNI--------KGVKPGTTASDNVD-------FVRTY 193
                | |   : |   || : :: || :           | |  ||    | |           :
Db       1572 DTDAVNFSQLK---SVSEAVDKGWTLTASGANGSKVVSGGTVDLKNTDGNLAISKSGDSN 1628

Qy        194 DTVEFLSADTK-------TTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKGKGEN 246
               | |  || | |         | || :    |   ::| :||    |          :|    |:
Db       1629 DVVFNLSKDFKVDEVTAGNTVVNTDGVKVG--SDVSLGAMGLFIANGPSVTASGFNAGDK 1686

Qy        247 -------GSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTG--------QADKFETV 291
                     |  :   :   |: : :|     | :|   |  |                  |
Db       1687 VISHVAVGMADTDAVNVSQLKQAVQSVTVKATRYYSTNDGGTQGGNYDGDGATGSKAIAA 1746

Qy        292 TSGTKVT-----------FASGNGTT-----ATVSKD----------DQG----NITVKY 321
               ||: :              ||| |:|     | | |             ||       : | ||
Db       1747 GVGTQASGEGAAAVGSGAAASGKGSTAIGRNAIASADGSVALGDGAKDGGRGAESYTGKY 1806

Qy        322 D---------VNVG-------------------DALNVNQL----QNSGWNLDS----- 343
                        |:||                     ||:|: ||      | |    |
Db       1807 SGVQNNTVGTVSVGDAAKGETRSISNVADAKEAMDAVNLRQLDAVAQKSNLQTDDMRHEI 1866

Qy        344 -------KAVAGSSGKVISGNVSPSKGKMDETVNINA---GNNIEITRN-----GKNIDI 388
                       |   | |   : |                    : :||    | |   ::      ||| ::
Db       1867 NNIEDVFKITKGDSASSVKG-----------MGVNAMAIGTNAAVSGTESVALGKNTNV 1914

Qy        389 ATSMTPQFSSVSLGAGADAPTLSVDDEGALNVGSKDANKPVRITNVAPGVKEGDVTNVAQ 448
              :         ::|::| |: |       |    :::|||   : :   ::||||  |  : |   ||:|
Db       1915 SAD-----NAVAIGNGSVA-----DRANSVSVGSGGSER--QVTNVAAGTADTDAVNVSQ 1962

Qy        449 LKGVAQNLNNRIDNVNGNAR----AGIAQAIATAGLVQAYLPGKSMMAIGGGTYLGEAGY 504
              |                | : || |       |:| ||||| | |||: |: | ::|  :| |
Db       1963 LNQGLITAKQYTDGMVGNLRRETSGGVAAAIATANLPQAYVQGRGMTSVGVSSYQGQSAI 2022

Qy        505 AIGYSSISAGGNWIIKGTASGNSRGHFGASASVGYQW 541
              |:| |::|  |:|: | : | |:| |  | |||||
Db       2023 AVGVSAVSESGHWVFKFSGSANTRSHVGVGAGVGYQW 2059
```

RESULT 5

```
A86036
probable adhesin Z5029 [imported] - Escherichia coli (strain O157:H7, substrain EDL933)
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C;Accession: A86036
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: A86036
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1588 &lt;STO&gt;
A;Cross-references: GB:AE005174; NID:g12518349; PIDN:AAG58749.1; GSPDB:GN00145; UWGP:
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: Z5029

  Query Match             13.8%;  Score 377;  DB 2;  Length 1588;
  Best Local Similarity   25.2%;  Pred. No. 2e-11;
  Matches  155;  Conservative   88;  Mismatches  267;  Indels  106;  Gaps   21;

Qy         18 GSIQASMEGSVELET--ISLSMTNDSKEFVDPYIVVTLKAGDNLKIKQNT---NENTNAS 72
              |:  || :|    : |    : |:: || : |:     :       |: |:|||    |:     :
Db        987 GAYNASHDGKASIITNVANGSISEDSTDAVNG----SQLNATNMMIEQNTQIINQLAGNT 1042

Qy         73 SFTYSLKKDLTGLINVET--EKLSF------GANGKKVNIISDTKGLNFAKETAGTNGDT 124
                 || ::::  |:  |  |   :  |:|          |      :    |  ||  :      |: |
Db       1043 DATY-IQENGAGINYVRTNDDGLAFNDASAQGVGATAIGYNSVAKGDSSVAIGQGSYSDV 1101

Qy        125 TVHLNGIGSTLTDMLLNTGA-TTNVTNDNV------TDDEKKRAASVKD------VLNAG 171
                :   |:::  ::   |:   |::|  : |        || |    |||  |  |: |       ::|
Db       1102 DTGIALGSSSVSSRVIAKGSRDTSITENGVVIGYDTTDGELLGALSIGDDGKYRQIIN-- 1159

Qy        172 WNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGK----KTEVKIGAKT 227
                   | |: | | |  ||            | | |   :  | :           : :||||
Db       1160 -----VADGSEAHDAVT-VRQLQNAIGAVATTPTKYFHANSTEEDSLAVGTDSLAMGAKT 1213

Qy        228 SVIKEK----------DGKLVTGKGKGE-------------NGSSTDEGEGLVTAKEVID 264
                 |  :|           |    :  |   |                 |||:|  |          :|
Db       1214 IVNGDKGIGIGYGAYVDANALNGIAIGSNAQVIHVNSIAIGNGSTTRGAQTNYTAYNMD 1273

Qy        265 AVNKAGWRMKTTTANGQ-------TGQADKFETVTSGTKVTFASGNGTTATVSKDDQ--- 314
                |    :         :|:||       | ||          ||| |   :   |  ::: |
Db       1274 APQNSVGEFSVGSADGQRQITNVAAGSADTDAVNVGQLKVTDAQVSQNTQSITNLDNRVT 1333

Qy        315 --GNITVKYDVNVGDAL----------NVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSK 362
                :      :     | :||        | :  :    | |  :  |  :   |  :|:
Db       1334 NLDSRVTNIENGIGDIVTTGSTKYFKTNTDGVDASAQGKDSVAIGSGSIAAADNSVALGT 1393

Qy        363 GKM---DETVNINAGNNIEITRN---GKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEG 416
              | :   : |::: :   |       |    |||    | ::       || : |    |          |
Db       1394 GSVATEENTISVGSSTNQRRITNVAAGKNATDAVNVAQLKSSEAGGVRYDTKADGSIDYS 1453

Qy        417 ALNVGSKDANKPVRITNVAPGVKEGDVTNVAQLKGVAQ-----------NLNNRIDNVNG 465
                : :|   :      ||:||: ||    || | ||||    |                 ::|::
Db       1454 NITLGGGNGG-TTRISNVSAGVNNNDVVNYAQLKQSVQETKQYTDQRMVEMDNKLSKTES 1512

Qy        466 NARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYLGEAGYAIGYSSISAGGNWIIKGTASG 525
                    ||| |:|  || ||| || || :|||||| ||:   |:| | :|| | |:  |    |
Db       1513 KLSGGIASAMAMTGLPQAYTPGASMASIGGGTYNGESAVALGVSMVSANGRWVYKLQGST 1572

Qy        526 NSRGHFGASASVGYQW 541
              ||:| : |:   ||
Db       1573 NSQGEYSAALGAGIQW 1588
```

RESULT 6

```
H91188
probable adhesin ECs4480 [similarity] - Escherichia coli (strain O157:H7, substrain R
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 02-Nov-2001
C;Accession: H91188
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and g
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: H91188
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1588 &lt;HAY&gt;
A;Cross-references: GB:BA000007; PIDN:BAB37903.1; PID:g13363955; GSPDB:GN00154
```

A;Experimental source: strain O157:H7, substrain RIMD 0509952
C;Genetics:
A;Gene: ECs4480

Query Match 13.8%; Score 377; DB 2; Length 1588;
Best Local Similarity 25.2%; Pred. No. 2e-11;
Matches 155; Conservative 88; Mismatches 267; Indels 106; Gaps 21;

```
Qy        18 GSIQASMEGSVELET--ISLSMTNDSKEFVDPYIVVTLKAGDNLKIKQNT---NENTNAS 72
Db       987 GAYNASHDGKASIITNVANGSISEDSTDAVNG----SQLNATNMMIEQNTQIINQLAGNT 1042

Qy        73 SFTYSLKKDLTGLINVET--EKLSF------GANGKKVNIISDTKGLNFAKETAGTNGDT 124
Db      1043 DATY-IQENGAGINYVRTNDDGLAFNDASAQGVGATAIGYNSVAKGDSSVAIGQGSYSDV 1101

Qy       125 TVHLNGIGSTLTDMLLNTGA-TTNVTNDNV------TDDEKKRAASVKD------VLNAG 171
Db      1102 DTGIALGSSSVSSRVIAKGSRDTSITENGVVIGYDTTDGELLGALSIGDDGKYRQIIN-- 1159

Qy       172 WNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGK----KTEVKIGAKT 227
Db      1160 -----VADGSEAHDAVT-VRQLQNAIGAVATTPTKYFHANSTEEDSLAVGTDSLAMGAKT 1213

Qy       228 SVIKEK----------DGKLVTGKGKGE-------------NGSSTDEGEGLVTAKEVID 264
Db      1214 IVNGDKGIGIGYGAYVDANALNGIAIGSNAQVIHVNSIAIGNGSTTTRGAQTNYTAYNMD 1273

Qy       265 AVNKAGWRMKTTTANGQ-------TGQADKFETVTSGTKVTFASGNGTTATVSKDDQ--- 314
Db      1274 APQNSVGEFSVGSADGQRQITNVAAGSADTDAVNVGQLKVTDAQVSQNTQSITNLDNRVT 1333

Qy       315 --GNITVKYDVNVGDAL----------NVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSK 362
Db      1334 NLDSRVTNIENGIGDIVTTGSTKYFKTNTDGVDASAQGKDSVAIGSGSIAAADNSVALGT 1393

Qy       363 GKM---DETVNINAGNNIEITRN---GKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEG 416
Db      1394 GSVATEENTISVGSSTNQRRITNVAAGKNATDAVNVAQLKSSEAGGVRYDTKADGSIDYS 1453

Qy       417 ALNVGSKDANKPVRITNVAPGVKEGDVTNVAQLKGVAQ----------NLNNRIDNVNG 465
Db      1454 NITLGGGNGG-TTRISNVSAGVNNNDVVNYAQLKQSVQETKQYTDQRMVEMDNKLSKTES 1512

Qy       466 NARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYLGEAGYAIGYSSISAGGNWIIKGTASG 525
Db      1513 KLSGGIASAMAMTGLPQAYTPGASMASIGGGTYNGESAVALGVSMVSANGRWVYKLQGST 1572

Qy       526 NSRGHFGASASVGYQW 541
Db      1573 NSQGEYSAALGAGIQW 1588
```

RESULT 7
A82615
surface protein XF1981 [imported] - Xylella fastidiosa (strain 9a5c)
C;Species: Xylella fastidiosa
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C;Accession: A82615
R;anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000
A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A;Reference number: A82515; MUID:20365717; PMID:10910347
A;Note: for a complete list of authors see reference number A59328 below
A;Accession: A82615
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1190 <SIM>
A;Cross-references: GB:AE004017; GB:AE003849; NID:g9107083; PIDN:AAF84783.1; GSPDB:GN001
A;Experimental source: strain 9a5c
R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H
as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.

submitted to GenBank, June 2000
A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Fr
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; La
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins
A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.
, F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri,
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawa
A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silv
M.; Tsuhako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.
A;Reference number: A59328
A;Contents: annotation
C;Genetics:
A;Gene: XF1981

Query Match 13.5%; Score 368.5; DB 2; Length 1190;
Best Local Similarity 23.9%; Pred. No. 3.8e-11;
Matches 153; Conservative 85; Mismatches 213; Indels 189; Gaps 25;

```
Qy        27 SVELE----TISLSMTNDSKEFVDPYIVVTLKAGDNLKIKQNTNENTNASSFTYSLKKDL 82
Db       615 SVDLKNSDGNLLITKTTDSND-------VTFNLATALKVDSLTTGNTAMTTDGVTVGSNV 667

Qy        83 ----TGLI-----NVETEKLSFGANGKKVNI---ISDTKGLNFA---------------- 114
Db       668 TLGSTGLVITDGPSVTSSGISAG-NQKITNVAAGTADTDAVNFSQLQAVSSTASKGWNLL 726

Qy       115 -----------------KETAG-------------------------TNGDTTVHLNGI 131
Db       727 ASGANSSNVAPGESVDLKNTDGNIVISKESGSNDVLFNLSSSLKLDKLTVGDTVMTTNGV 786

Qy       132 GSTLTDMLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAG-WNIKGVKPGTTASD----- 185
Db       787 -------TVGSGVTLGSMGLVITDGPSVTSSGI----NAGSQKITNVAAGTADTDAVNLS 835

Qy       186 ------------NVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEK 233
Db       836 QLNTAMAGSGAKSVHYYSTYD------GGTQGGNYNGDGATGTRSIAVGVGTLASA---- 885

Qy       234 DGKLVTGKGKGENG-SSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVT 292
Db       886 EGATAVGSGAAASGKGSTAIGRNAVASADGSVALGD-GAKDGARGAESYTGKYSGLQNNT 944

Qy       293 SGTKVTFASGNGTTATVSKDDQGNITVKYDVNVGDALNVNQL----QNSGWNLDSKAVAG 348
Db       945 VGTVSVGDASKGETRTVS-----NVADAKEAT--DAVNLRQLDRVAQDANRYVDNKIESL 997

Qy       349 SSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADA- 407
Db       998 SEGQTF--------------VKVNSLNN-----------SATPIAAGVDATAIGVGATAS 1032

Qy       408 ----------PTLSVDDEGAL------------NVGSKDANKPVRITNVAPGVKEGDVTN 445
Db      1033 GADSIAMGNKASASADNAVAIGNHSVADRANTVSVGSAGSER--QVTNVAAGTADTDAVN 1090

Qy       446 VAQLKGVAQNLNNRIDNVNGNAR----AGIAQAIATAGLVQAYLPGKSMMAIGGGTYLGE 501
Db      1091 VSQLNQGLITAKQYTDGVVGSLRRDTDGGVAAAIATANLPQAYIPGRGMTSVGVSSYRGQ 1150

Qy       502 AGYAIGYSSISAGGNWIIKGTASGNSRGHFGASASVGYQW 541
Db      1151 SAIAVGVSSVSESGRWVFKFSGSANTRSQVGIGAGVGYQW 1190
```

RESULT 8
AC0976
probable autotransporter sapB [imported] - Salmonella enterica subsp. enterica serova
C;Species: Salmonella enterica subsp. enterica serovar Typhi
A;Note: this species has also been called Salmonella typhi
C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C;Accession: AC0976
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Church
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farr
, S.; Moule, S.; O'Gaora, P.

Nature 413, 848-852, 2001
A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A;Reference number: AB0502; MUID:21534947; PMID:11677608
A;Accession: AC0976
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1107 <PAR>
A;Cross-references: GB:AL513382; PIDN:CAD03303.1; PID:g16504923; GSPDB:GN00176
C;Genetics:
A;Gene: sapB

Query Match 13.3%; Score 363; DB 2; Length 1107;
Best Local Similarity 22.1%; Pred. No. 6.6e-11;
Matches 167; Conservative 95; Mismatches 239; Indels 256; Gaps 26;

```
Qy         29 ELETISLSMTNDSKEF-----VDPYIVVTLKAGDNLKIKQNTNENTNASSFTYSLKKDLT 83
Db        363 ELSEESTDAVNGSQLFETNEKVDQN--TTDIAANTTNITQNSTAIENLNTSVSDINTSIT 420

Qy         84 GLI------NVETEKLSFGANGKKVNIISDTKGLNFAKETAGTNG----DTTVHLNGIGS 133
Db        421 GLTDNALLWDEDTGAFSANHGGSTSKITNVAAGALSEDSTDAVNGSQLYETNQKVDQNTS 480

Qy        134 TLTDMLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFV--- 190
Db        481 AIAD--INT-SITNLGTDALSWDDEEGAFSASHGTSGTNKITNVAAGEIASDSTDAINGS 537

Qy        191 RTYDT----------VEFLSADTKTTTVN--------VESKDNGKKTEVKIGAKTSVIKE 232
Db        538 QLYETNMLISQYNESISQLAGDTSETYITENGTGVKYIRTNDNGLEGQDAYATGNGATAV 597

Qy        233 KDGKLVTGKG---KGENGSSTDEG-------------------------EGLV-----TA 259
Db        598 GYDAVASGAGCLALGQNSSSSIEGSIALGSGSTSNRAITTGIRETSATSDGVVIGYNTTD 657

Qy        260 KEVIDAVNKA----GWRMKTTTANGQTGQ-------------------ADKFETVTSGTKV 297
Db        658 RELLGALSLGTDGESYRQITNVADGSEAQDAVTVRQLQNAIGAVTTTPTKYYHANSTEED 717

Qy        298 TFASGNGTTATVSKDDQGNITVKYDVNVGDALNVNQLQN--SGWNLDSKAVAGSSGKVIS 355
Db        718 SLAVGTDSLAMGAK-----TIVNADAGIGIGLNTLVMADAINGIAIGSNARANHANSIAM 772

Qy        356 GNVSP------------------------SKGKMD---ETVNINAGN------------ 375
Db        773 GNGSQTTRGAQTDYTAYNMDTPQNSVGEFSVGSEDGQRQITNVAAGSADTDAVNVGQLKV 832

Qy        376 -NIEITRNGKNI----------------------DIATSMTPQF--------------S 397
Db        833 TDAQVSRNTQSITNLNTQVSNLDTRVTNIENGIGDIVTTGSTKYFKTNTDGADANAQGAD 892

Qy        398 SVSLGAGADAP--------TLSVDDE-GALNVGSKDANKPVRITNVAPGVKEGDVTNVAQ 448
Db        893 SVAIGSGSIAAAENSVALGTNSVADEANTVSVGSSTQQR--RITNVAAGVNNTDAVNVAQ 950

Qy        449 LK-------------------------------------------------GVAQ------- 454
Db        951 LKASEAGSVRYETNADGSVNYSVLNLGDGSGGTTRIGNVSAAVNDTDAVNYAQLKRSVEE 1010

Qy        455 ----------NLNNRIDNVNGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYLGEAGY 504
Db       1011 ANTYTDQKMGEMNSKIKGIENKMSGGIASAMAMAGLPQAYAPGANMTSIAGGTFNGESAV 1070

Qy        505 AIGYSSISAGGNWIIKGTASGNSRGHFGASASVGYQW 541
Db       1071 AIGVSMVSESGGWVYKLQGTSNSQGDYSAAIGAGFQW 1107
```

RESULT 9
AH0110
probable surface protein (partial) YPO0902 [imported] - Yersinia pestis (strain CO92)
C;Species: Yersinia pestis
C;Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 02-Nov-2001
C;Accession: AH0110
R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrel
Nature 413, 523-527, 2001
A;Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A;Reference number: AB0001; MUID:21470413; PMID:11586360
A;Accession: AH0110
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-658 <KUR>
A;Cross-references: GB:AL590842; PIDN:CAC89747.1; PID:g15978974; GSPDB:GN00175
C;Genetics:
A;Gene: YPO0902

Query Match 11.8%; Score 321.5; DB 2; Length 658;
Best Local Similarity 23.0%; Pred. No. 4.2e-09;
Matches 138; Conservative 71; Mismatches 214; Indels 177; Gaps 21;

```
Qy         64 NTNENTNASSFTYSLKKDLTGLINVE----TEKLSFGANGKKV---NIISDTKGL---NF 113
Db        114 STSTDAVVGSQLYNLVQDGTRYFHANSVNPTDSLASGLETIAVGPATVVSGDNGVGIGNT 173

Qy        114 AKETAGTNGDTTVHLNGIGSTLT----------------DMLLNTGATTNVTNDNVTDD 156
Db        174 ALVGAAATGGIAI---GFGTQVTAAGATAIGSAAQAQGAQSLALGAGAVTSQANSIALG- 229

Qy        157 EKKRAASVKDV-------------------LNAGW-----NIKGVKPGTTASDNVDFVR 191
Db        230 ----AASINTVGAQSSYSAYALTAPQASVGELGIGTALGNRKITGVAAGSASSDAVNVAQ 285

Qy        192 --------TYDTVEFLSADTKTTTV--NVESKDNG----------KKTEVKIGAKTSVI 230
Db        286 LTAVGDQVQQNTANITSLGGRVTTIEGSMASIANGGGVKYFHANSTQPDSVASGTNSVAI 345

Qy        231 -------------------------KEKDGKLVTGKGKGENGSSTDEGEGLVTAKEVID 264
Db        346 GPASLASGNAALASGAGAVAIGDGAAASADGSVAIGQGSGDNGRGVE------------- 392

Qy        265 AVNKAGWRMKTTTANGQTGQADKFETVTSGTKVTFASGNGTTATVSKDDQGNITVKYDVN 324
Db        393 ---------------NYIGKYSNASNTSSGTVSVGNTATGETRTVSNVADG-------LQ 430

Qy        325 VGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNI-EITRNG 383
Db        431 ATDAVNLRQLDG----------IAASIVVVENNVSGLQNGTDGMFQVNNSSGLAKPSATG 480

Qy        384 KNIDI--ATSMTPQFSSVSLGAGADAPTLS---------VDDEGALNVGSKDANKPVRIT 432
Db        481 ANSATGGAGSVASGNNSTAFGSGAKATAANSAALGANSVADRANSVSVGSVGNER--QIT 538

Qy        433 NVAPGVKEGDVTNVAQLKGVAQN----LNNRIDNVNGNAR-------AGIAQAIATAGLV 481
Db        539 NVAPATQGTDAVNFDQLKSISNQTNAYTNQRYSELKQDLRKQNSVLSAGIASAMSMASLT 598

Qy        482 QAYLPGKSMMAIGGGTYLGEAGYAIGYSSISAGGNWIIKGTASGNSRGHFGASASVGYQW 541
Db        599 QPYTSGSSMTTIGAASYRGQSALSLGVSSISDSGRWVSKLQASSNTQGDFGIGVGVGYQW 658
```

RESULT 10
A43855
high-molecular-weight surface-exposed protein HMW1 - Haemophilus influenzae
C;Species: Haemophilus influenzae
C;Date: 10-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 08-Oct-1999
C;Accession: A43855
R;Barenkamp, S.J.; Leininger, E.
Infect. Immun. 60, 1302-1313, 1992
A;Title: Cloning, expression, and DNA sequence analysis of genes encoding nontypeable
detella pertussis.
A;Reference number: A43855; MUID:92192797; PMID:1548058
A;Accession: A43855

```
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1536 <BAR>
A;Cross-references: GB:U08876; GB:M84616; NID:g475770; PIDN:AAA20527.1; PID:g475771
A;Note: sequence inconsistent with the nucleotide translation
A;Note: sequence extracted from NCBI backbone (NCBIN:89235, NCBIP:89239)

  Query Match             8.4%;  Score 231;  DB 2;  Length 1536;
  Best Local Similarity  23.4%;  Pred. No. 0.00049;
  Matches  142;  Conservative   90;  Mismatches  215;  Indels  160;  Gaps   33;

Qy         18 GSIQASMEGSVELETISLSMTNDSKEFVDPYIVVTLK-----AGDNLKIKQNTNENTNAS 72
              |:|  |  |||:   ::        |      | :|:  |       | :|: | : :  |
Db        739 GNISVSGGGSVDFTLLA-----SSSNVQTPGVVINSKYFNVSTGSSLRFKTSGSTKTG-- 791

Qy         73 SFTYSLKKDLT-----GLINVETEKLSFGANGK----KVNIISDTKGLNFAKETAGT--N 121
                 :|::||||      | | :    : : |   ||     | ||   :    : |    | |
Db        792 ---FSIEKDLTLNATGGNITLLQVEGTDGMIGKGIVAKKNITFEGGNITFGSRKAVTEIE 848

Qy        122 GDTTVHLNG----IGSTL----------TDMLLNTGATT---NVTN--DNVTDDEKKRAA 162
              |: |:: |         |||              |:::|:|  |     |: |    |:| :     |
Db        849 GNVTINNNANVTLIGSDFDNHQKPLTIKKDVIINSGNLTAGGNIVNIAGNLTVESN---A 905

Qy        163 SVKDVLNAGWNIKGV-----------KPGTTASD---------NVDFVRTYDTVEFLSA 201
              : |  : |  :|: |:           | |      |              :    || |:    :
Db        906 NFKAITNFTFNVGGLFDNKGNSNISIAKGGARFKDIDNSKNLSITTNSSSTYRTIISGNI 965

Qy        202 DTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKGK-------------GENG 247
                |   :|: ::  |  ||::||   |    :|:| |      |                  |||
Db        966 TNKNGDLNITNE--GSDTEMQIGGDVS---QKEGNLTISSDKINITKQITIKAGVDGENS 1020

Qy        248 SSTDEGEGLVTAK----EVIDAVNKAGWRMKTTTAN-------GQTGQADKFETVTSGTK 296
               |         :| |    ::   :|`:|:     ||          | |  ||      |:  |
Db       1021 DSDATNNANLTIKTKELKLTQDLNISGFNKAEITAKDGSDLTIGNTNSADG----TNAKK 1076

Qy        297 VTF---------ASGNGTTATVSKDDQGNITVKYDVNVGDALNVNQLQNSGWNLDSKAVA 347
              |||          | |:  |       :   |:       :  |   |:  :       |:|   :|:| |
Db       1077 VTFNQVKDSKISADGHKVTLHSKVETSGS-----NNNTEDSSD----NNAGLTIDAKNV- 1126

Qy        348 GSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIAT---SMTPQFSSVSLGAG 404
                   :: |::  |      |:|:| :      |:  |  |:  |      :| |  |:  :
Db       1127 -----TVNNNITSHK-----AVSISATSGEITTKTGTTINATTGNVEITAQTGSILGGIE 1176

Qy        405 ADAPTLSVD-DEGALNVGSKDANKPVRITNVAPGVKEGDVTNVA--QLKGV-AQNLNNRI 460
               : : ::::    ||||  | :    |         |          | :| :|     :||  :     :::
Db       1177 SSSGSVTLTATEGALAVSNISGN------TVTVTANSGALTTLAGSTIKGTESVTTSSQS 1230

Qy        461 DNVNGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYLGEAGYAIGYSSISAGGNWIIK 520
               :: |     |  :  ||` |         :|   |    |   |||          || |    |
Db       1231 GDIGGTISGGTVEVKATESLTT-----QSNSKIKATT--GEANVT------SATGT--IG 1275

Qy        521 GTASGNS 527
              || |||:
Db       1276 GTISGNT 1282
```

RESULT 11
A56143
surface-array protein homolog sapA2 - Campylobacter fetus (strain 82-40 LP3)
C;Species: Campylobacter fetus
C;Date: 03-Nov-1995 #sequence_revision 03-Nov-1995 #text_change 08-Oct-1999
C;Accession: A56143
R;Dworkin, J.; Tummuru, M.K.R.; Blaser, M.J.
J. Bacteriol. 177, 1734-1741, 1995
A;Title: A lipopolysaccharide-binding domain of the Campylobacter fetus S-layer protein
A;Reference number: A56143; MUID:95204338; PMID:7896695
A;Accession: A56143
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1109 <DWO>
A;Cross-references: GB:S76860; NID:g913763; PIDN:AAB33871.1; PID:g913764
A;Note: this gene appeared to be silent in strain 82-40 LP3
C;Genetics:
A;Gene: sapA2

```
  Query Match             8.2%;  Score 225;  DB 2;  Length 1109;
  Best Local Similarity  25.0%;  Pred. No. 0.00064;
  Matches  152;  Conservative   78;  Mismatches  221;  Indels  158;  Gaps   35;

Qy          4 DEEEELESVQRSVVGSIQASMEGSVE-LETISLSMTNDSKEFVDPYIV-VTLKAGDNLKI 61
              |   ::::: : ||:  ||:  : : :||::|: |   |: ||    :  :::|   || :
Db        330 DGSADVQNLKVNGVGAKGASVAITADKIETLNLN-TTGSQSFVSADVASISVKGNANLSL 388

Qy         62 KQNTNENT-NASSFTYSLKKDLTGLINVETEKLSFGANGKKVNIISDTKGLNFAKETAGT 120
                     | :||||  :|   ||:.   :| : |    | ||       | |    :| | :  |
Db        389 ATGAKTTTLDASSFGGALDADLSTSASVTSIK---GGNGNDKITIKDV-AVNVAID-GGA 443

Qy        121 NGDTTVHLNGIGSTLTDMLLNTGATTNVTNDNVTDD---EKKRAASVKDVLNAGWNIKGV 177
                 |  |          ||   | |        || |   | |         |:| || ::       : | :
Db        444 GNDELVIKGSTADTLQPTLTN---IEKVTIDGNTKDLTLSLKKAQSVTEL-----SFKNI 495

Qy        178 -KPGTTASDNVDFVRTYDTVEFL--SADTKTTTVNVES---------KDNGKKTEVKIGA 225
                |  | :: ||       :||  |  :|  |  |:| ||                | |       : |
Db        496 AKTVTESNGNV------ETVNILANNATDKAVTINDESLKTINFSDVDDKG----ASVAA 545

Qy        226 KTSVIKEKDGKLVTGKGK----------GENGSSTDEGEGLVTAKEVID----------- 264
                |  :: :|  :|           |           | :  |      :  ||: :
Db        546 KGKIVADKATELTINSNKVTLASDAVVQAANATKID----INAAKDTVGLTLGGVAKLTD 601

Qy        265 -AVNKAGWRMKTTTANG-----------QTGQADKFETVTS----------GTKVTFASG 302
                 ||  |         | ||                   |  |    | ||          |     |
Db        602 LTVNNKG-AFALTGANATDLDSVKNLSVNTEGAFSIATATSLKNLNNLSLNGVSADLNSV 660

Qy        303 NGTTATVSKDDQ----------GNITVKYDV-----NVGDALNVNQLQNSGWNLDSKAVA 347
               |  |||::  :             |  |  |  ||       ||| ||  :  : :|  |  |  ::
Db        661 NVGTATLASLEANINVSGEFKLGTTTAKGDVDFNIENVG-ALTLGAITSSTGNA-SVIIS 718

Qy        348 GSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADA 407
                ::|  |   | ||  ::| :        :||||  :                    ::::||
Db        719 SATGNVTLGAVSATQGNL----TLNAGNTL--------------------GNITIGA-LKG 754

Qy        408 PTLSVDDEGALNVGSKDANKPVRITNVAPGVKEGDVT--------NVAQLKGVA--QNLN 457
                :|||  | |    : |||  | ||:            :||           || ::   |    :||
Db        755 DIVSVDLGGVLGTINSDANNKVSITS-------NEVTYVGSEISKNVVEITAAAGGTDLN 807

Qy        458 NRIDNVNGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYLGEAGYAIGYSSISAGGNW 517
              :     |  |  |  |  |   |  |   | |             : |||    |      ||:  |    |
Db        808 AQ---VIGGAAADDALTIIGKGDTQTITASGD---LSGGTLTLTLTEATKLSSLDISG-- 859

Qy        518 IIKGTASGN 526
               :||  :||
Db        860 -VKG-ITGN 866
```

RESULT 12
C82672
surface-exposed outer membrane protein XF1516 [imported] - Xylella fastidiosa (strain
C;Species: Xylella fastidiosa
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C;Accession: C82672
R;anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Seq
Nature 406, 151-157, 2000
A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A;Reference number: A82515; MUID:20365717; PMID:10910347
A;Note: for a complete list of authors see reference number A59328 below
A;Accession: C82672
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1004 <SIM>
A;Cross-references: GB:AE003981; GB:AE003849; NID:g9106543; PIDN:AAF84325.1; GSPDB:GN
A;Experimental source: strain 9a5c
R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer
as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Fr

J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigr
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E
A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
, F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
M.; Tsuhako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A;Reference number: A59328
A;Contents: annotation
C;Genetics:
A;Gene: XF1516

Query Match 8.1%; Score 220.5; DB 2; Length 1004;
Best Local Similarity 23.0%; Pred. No. 0.00096;
Matches 142; Conservative 76; Mismatches 236; Indels 163; Gaps 29;

```
Qy         34 SLSMTNDSKEFVDPYIVVTLKAGDNLKIKQNTNENTN-ASSFTYSLKKDLTGLINVETEK 92
Db        114 SIAIGYFSKAFAPNAIAL----GYNSSVTQSANNGVALGSNSTVS---------GVNSVA 160

Qy         93 LSFGANGKKVNIISDTKGLNFAKETAGTNGDTTVHLNGIGSTLTDMLLNTGATTNVTNDN 152
Db        161 LGAGSMASELNVISVGGG-------DGVTGPAVRRIVNVG----DGIGNNDAVNKSQLDG 209

Qy        153 VTDDEKKRAASVKDVL------------NAGWNIKGVKPGTTA-SDNVDFVRTYDTVEF- 198
Db        210 VTASVNDVAASVKTIALTNQVTGSSVASASGKESTAIGSGAQAVADN--------TVAFG 261

Qy        199 ------------LSADT-----KTTTVNVESKD--------------NGKKTEVKIGAK 226
Db        262 GRAIANAVGASALGFDSHAKGINSTTVGTQSVSLGQGGVSLGYNSFVGEGSFNGLALGSN 321

Qy        227 TSVIKEKDGKLVTGKG---KGENGSSTDEGEGL--VTAKEVI---------DAVNKAGWR 272
Db        322 SLVLLQGVDSVALGSGSMASEPNVVSVGSGDGLRGPAVRRIVNVGDGIGNNDAVNKSQLD 381

Qy        273 MKTTTANGQTGQADKFETV--TSGTKVTFASGNGTTAT-VSKDDQGNITVKYDVNVGDAL 329
Db        382 GVTASVNDVVASVKNIAGAIQITGSGVASVSGQDSTAAGASAQAAGDSSIA----LGARS 437

Qy        330 NVNQLQNSGWNLD-------SKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRN 382
Db        438 RANAIGSSALGVDGHALGANSTALGGQSTAISEGGTSLG---YNSFVGQSATNGIALGSN 494

Qy        383 GKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGALNVGSKD-ANKPV--RITNVAPGVK 439
Db        495 --------AIVSGVNSVALGAGSVASELNV-----ISVGGGDGVTGPAVRRIVNVGDGIG 541

Qy        440 EGDVTNVAQLKGVAQNLNN---RIDNVNGNAR---AGIAQAI---ATAGLVQAYLPGKSM 490
Db        542 NNDAVNKSQLDGVTASVNDVAASVKKIVGTIQITGSGVASAIGKDSTATGASAQAVGDSS 601

Qy        491 MAIG----------------------GGTYLGEAGYAIGYSSISAGGNWIIK-----GTA 523
Db        602 VALGTRATANAIGSSVLGVDSRARGINSTALGRQSNAIGDGSVSLGFNSFVRQSGEHGVA 661

Qy        524 SGNSRGHFGA-SASVGY 539
Db        662 LGTDAGVSGKDSIALGY 678
```

RESULT 13
AF0394
probable adhesin hmwA [imported] - Yersinia pestis (strain CO92)
C;Species: Yersinia pestis
C;Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 02-Nov-2001
C;Accession: AF0394
R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,
Nature 413, 523-527, 2001
A;Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A;Reference number: AB0001; MUID:21470413; PMID:11586360
A;Accession: AF0394
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1910 <KUR>
A;Cross-references: GB:AL590842; PIDN:CAC92482.1; PID:g15981183; GSPDB:GN00175
C;Genetics:
A;Gene: hmwA

Query Match 7.8%; Score 214; DB 2; Length 1910;
Best Local Similarity 22.1%; Pred. No. 0.0047;
Matches 141; Conservative 81; Mismatches 219; Indels 198; Gaps 31;

```
Qy         27 SVELETI--SLSMTNDSKEFVDPYIVVTLKAGDNLKIKQ----------NTNENT----N 70
Db       1166 NVELNATAGNVSIYAETKTALSTSLNAVLSLGGNNSIKAQNGWLIGKAFNTTQGAGIGFR 1225

Qy         71 ASSFTYSLKKDLTGLINVETEKLS--------FGAN------GKKVNIISDTKGLNFAKE 116
Db       1226 ANS---SLSVDGNIILKGETEGVGATRKGIDFYGANTLNIIKGSQLSLLGENKG---AQD 1279

Qy        117 TAGTNG-------DTTVHLNGI----GSTLTDMLLNTGATTNVTNDNVTDDEKKRAASVK 165
Db       1280 TAGGNGISYTSLAKLTVNNNGSLKMEGRSTSGTGINFPSSNNTLVFNGDGDTLIKGSS-- 1337

Qy        166 DVLNAGWNIKGVKPGTTASDNVDFVRTYDT-VEFLSADTKTTTVNVESKDNGKKTE---V 221
Db       1338 -VAGTGAAISGVVNNSTGPMTIEGISTDGAGVHLFSAEHRIDRINV----TGSSTHAEGL 1392

Qy        222 KIGAKTSVIKEKDGKLVTGKGKGENGSSTD----EGEGLVTAKEVIDAVNKAGWRMKTTT 277
Db       1393 RISGNAAIVD------TTLTGKSINGSGVKIDSLPGSSVVT-RSVLD----------NAT 1435

Qy        278 ANGQTGQADKFETVTSGTKVTFASGNGTTATVSKDDQGNITVKYDVNVGDALNV------ 331
Db       1436 LNGSSSSGKGVEITSDINGIHHSSINGTTTGTG----------YGIDIGENSNVTGTSEA 1485

Qy        332 ----------------------NQLQNSGWNLDSKAVAG----------SSGKVISGNVS 359
Db       1486 DLLILQGVATTGTGTGIKLNGNNDLSNT--SLNSSAVDGIALDITGPLANQGNVILNGTA 1543

Qy        360 PSKG--------KMDETVNINAGNNIEITRN---------------------------- 382
Db       1544 SGSGIGAQVNGSLSDSVVNGTSTNGIGVQINGSLKNSRINGISANGSGVKIDGESTLDNA 1603

Qy        383 --------GKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGALNVGSKDANKPVRITNV 434
Db       1604 TLNGNSTEGKGVDLAANLSGNHGSVVHGDTVHGTGIDVGKDVTLSGGGTD--EPLTVSGN 1661

Qy        435 APGVKEGDVTNVAQLKGVAQNLNNRIDN--VNGNARAGIAQAIATAGLVQAYLPGKSMMA 492
Db       1662 ASGEKGTGV----QLGG-----NNTLDNTTLSGNATDGHGVEI------NSRLINNGNTT 1706

Qy        493 IGGGTYLGEAGYAIGYSSISAGGNWIIKGTASGNSRGHF 531
Db       1707 INGKT--SDDGHGVHINGAISGGE--ING-HSDNSHGVF 1740
```

RESULT 14
T31105
hypothetical protein 2 - Haemophilus ducreyi
C;Species: Haemophilus ducreyi
C;Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999
C;Accession: T31105
R;Ward, C.K.; Lumbley, S.R.; Latimer, J.L.; Cope, L.D.; Hansen, E.J.
J. Bacteriol. 180, 6013-6022, 1998
A;Title: Haemophilus ducreyi secretes a filamentous hemagglutinin-like protein.
A;Reference number: Z20984; MUID:99030326; PMID:9811662
A;Accession: T31105
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-4919 <WAR>
A;Cross-references: EMBL:AF057696; NID:g3929021; PID:g3929023; PIDN:AAC79761.1
C;Genetics:

A;Gene: lspA2

Query Match 7.8%; Score 212.5; DB 2; Length 4919;
Best Local Similarity 21.8%; Pred. No. 0.019;
Matches 151; Conservative 85; Mismatches 248; Indels 209; Gaps 32;

```
Qy         17 VGSIQASME-GSVELETISLSMTNDSKEFVDPYIVVTLKAGD------------------ 57
Db        429 ISALNLTLENATVSANNLSFRVTNDTK--LNNLSKVSARAADLQSGNLNLDKASVLAHKL 486

Qy         58 NLKIKQNTNENTNASSFTYSLKKDLTGLINVETEKLSFGANGKKVNIISDTKGLNFAKET 117
Db        487 TLNISNDVSLNNQSKLSANNLKIKKVRDLNLNNSELS--ANNLTLNTSNNITLKNKSKFT 544

Qy        118 AGT---NGDTTVHLNGIGSTLTDMLLNTGATTNVT----------------NDNVTDDEK 158
Db        545 AGNMTLNVTNNVTLNN-DSELAANNLTLNVTKNVTLNDASKLSANKLDLNVTDNVTLNSK 603

Qy        159 KRAAS-------VKDV--------------LNAGWNIKGVKPGTTASDNVDF----VRTY 193
Db        604 STLSAGELTFKKVKNVTLNNDSELAANNLSLNASHNVTLNNKSKLSAQKADIKAVNLTLN 663

Qy        194 DTVEFL--SADTKTTTV-----------NVESKDNGKKTEVKIGAKTSVIKEKDGKLVTG 240
Db        664 DTTELTAKNLDINSTTITNNGTIAGIFANITTEKLNNKEKALILAEQNLNFTVNGSHYEN 723

Qy        241 KGK-----------GENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFE 289
Db        724 KGDIVSKDKATVTFSKNSDFTSNGSKLVNAQN----------QLKVNVNNFTISQGDDIT 773

Qy        290 ---TVTSGTKVTFA-SGNGTTA-TVSKDDQGNITVKYDVNVGDALNV---NQLQNSG--- 338
Db        774 LIGNVTLNASGTFTNSGNLTTVKTLDVGDIQNFTNKGNLTVGEDLHIKSKTKITNDGKLI 833

Qy        339 ----WNLDSKAVAGSSGKVI---------SGN---------------VSPSKGKMDETV 369
Db        834 SIKNLNISSEADFINNGTLLGIEALKIATKGNFTNKEKAILASNSLLDISVAEGK--KTF 891

Qy        370 N---INAGNNIEITRNGK--NIDIATSMTPQFSSVSLGAGADAPTLSVDDEGALNVGSK- 423
Db        892 NNGTIESGKNLNITNTGAFLNVDNATIR-------SFGVLNITSTGNVSNNGTLISNERL 944

Qy        424 DANKPVRITNVAPG----------VKEGDVTNVAQLKGVAQ-NLNNRIDNVNGNARAGI 471
Db        945 NITSAANFTNESNGTVMSNGLLNIIAKQGNITNKNLIASRQQLNLTAVADNITND--SNI 1002

Qy        472 AQAIATAGLVQAYLPGKSMMAIGGGTYLGEAGYAIGYSSISAGGNWIIK---------- 520
Db       1003 SNKIAVLHSL-------------GNISLNSKDQVYNLGEIYAGNNISVKAHQLKNDVKLM 1049

Qy        521 ---------GTAS-----GNSRGHFGASASVGY 539
Db       1050 GDITTKTKEGQASYKLYQASNGGHFGNDGSSGY 1082
```

RESULT 15
A41477
190K surface antigen precursor - Rickettsia rickettsii
C;Species: Rickettsia rickettsii
C;Date: 16-Sep-1992 #sequence_revision 16-Sep-1992 #text_change 08-Oct-1999
C;Accession: A41477
R;Anderson, B.E.; McDonald, G.A.; Jones, D.C.; Regnery, R.L.
Infect. Immun. 58, 2760-2769, 1990
A;Title: A protective protein antigen of Rickettsia rickettsii has tandemly repeated, ne
A;Reference number: A41477; MUID:90354033; PMID:2117568
A;Accession: A41477
A;Molecule type: DNA
A;Residues: 1-2249 <AND>
A;Cross-references: GB:M31227; NID:g152465; PIDN:AAA26380.1; PID:g152466
A;Note: the authors translated the codon GAT for residue 430 as Gly, and CAA for residue
C;Keywords: surface antigen; tandem repeat
F;1-20/Domain: signal sequence (uncleaved) #status predicted <SIGU>

Query Match 7.7%; Score 210; DB 2; Length 2249;
Best Local Similarity 23.8%; Pred. No. 0.0092;
Matches 139; Conservative 68; Mismatches 255; Indels 122; Gaps 27;

```
Qy          9 LESVQRSVVGSIQASMEGSVELETISLSMTNDSKEFVDPYIVVTLKAGDNLKIKQNTNEN 68
Db        114 LNITQNTVVGSI--ITKGNL--------------------LPVTLNAGKSLTLNGN---N 148

Qy         69 TNASSFTYSLKKD-LTGLINVETEKLSFGANG---------KKVNIIS--DTKGLNFAKE 116
Db        149 AVAANHGFDAPADNYTGLGNIALG----GANAALIIQSAAPSKITLAGNIDGGGIITVKT 204

Qy        117 TAGTNGDTTVHLNGIGSTLTDMLLNTGATTNVTNDNV---TDDEKKRAASVKDVLNAGWN 173
Db        205 DAAINGT-------IGNTNALATVNVGAGTATLGGAVIKATTTKLTNAASVLTLTNANAV 257

Qy        174 IKGVKPGTTASDNVDFVRTYDTVEFLSADTKTT----TVNVESKD---NG---KKTEVKI 223
Db        258 LTGAIDNTTGGDNVGVLNLNGALSQVTGDIGNTNSLATISVGAGTATLGGAVIKATTTKL 317

Qy        224 GAKTSVIKEKDGKLVTG----KGKGENGSSTDEGEGLVTAK----EVIDAVNKAGWRMKT 275
Db        318 TDAASAVKFTNPVVVTGAIDNTGNANNGIVTFTGNSTVTGNVGNTNALATVNVGAGLLQ- 376

Qy        276 TTANGQTGQADKFETVTSGTKVTFASGNGTTATVSKDDQGNIT---VKYDVN---VGDAL 329
Db        377 --VQGGVVKANTINLTDNASAVTFT--NPVVVTGAIDNTGNANNGIVTFTGNSTVTGDIG 432

Qy        330 NVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINA--GNNIEITRNGKNID 387
Db        433 NTNALATVNVGAGTATLGGAVIKATTTKLTNAASVLTLT-NANAVLTGAIDNTTGGDNVG 491

Qy        388 IAT------------SMTPQFSSVSLGAGADAPTLSVDDEGALNVGSKDANKPVRITNVA 435
Db        492 VLNLNGALSQVTGNIGNTNSLATISVGAG--TATLGGAVIKATTTKLTDAASAVKFTN-- 547

Qy        436 PGVKEGDVTNVAQLKGVAQNLNNRIDNVNGNARA----GIAQAIATAGLVQAYLPGKSMM 491
Db        548 PVVVTGAIDNTG-------NANNGIVTFTGNSTVTGDIGNTNSLATISV------GAGTA 594

Qy        492 AIGGG------TYLGEAGYAIGYSSISAGGNWIIKGTASGNSRG 529
Db        595 TLGGAVIKATTTKLTNAASVLTLTNANAVLTGAIDNTTGGDNVG 638
```

Search completed: October 6, 2003, 09:33:32
Job time : 19.4863 secs

OM protein - protein search, using sw model

Run on: October 6, 2003, 09:06:20 ; Search time 8.35443 Seconds
(without alignments)
3045.266 Million cell updates/sec

Title: US-09-771-382-34
Perfect score: 2735
Sequence: 1 TDEDEEEELESVQRSVVGSI.....TASGNSRGHFGASASVGYQW 541

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	210	7.7	2249	1	OMPA_RICRI	P15921 rickettsia
2	206.5	7.6	1655	1	OMPB_RICCN	Q9kka3 r outer mem
3	201	7.3	1654	1	OMPB_RICRI	Q53047 r outer mem
4	198	7.2	2003	1	YDBA_ECOLI	P33666 escherichia
5	197	7.2	1577	1	HLYA_PROMI	P16466 proteus mir
6	189.5	6.9	2021	1	OMPA_RICCN	Q52657 rickettsia
7	186	6.8	1656	1	OMPB_RICJA	O06653 r outer mem
8	184.5	6.7	1953	1	BIGA_SALTY	P25927 salmonella
9	180	6.6	933	1	SLAP_CAMFE	P35827 campylobact
10	179.5	6.6	737	1	ALYS_ENTFA	P37710 enterococcu
11	179	6.5	1608	1	HLYA_SERMA	P15320 serratia ma
12	179	6.5	1645	1	OMPB_RICTY	P96989 r outer mem
13	177	6.5	1039	1	AG43_ECOLI	P39180 escherichia
14	175.5	6.4	1325	1	YDEK_ECOLI	P32051 escherichia
15	173	6.3	1567	1	ICEN_XANCT	P18127 xanthomonas
16	172	6.3	1025	1	SLAP_CAUCR	P35828 caulobacter
17	172	6.3	1300	1	120K_RICRI	P14914 rickettsia
18	170.5	6.2	1153	1	PVDB_PLAKN	P50493 plasmodium
19	170	6.2	2660	1	YEEJ_ECO57	Q8x8v7 escherichia
20	168.5	6.2	1569	1	YPJA_ECOLI	P52143 escherichia
21	166.5	6.1	1643	1	OMPB_RICPR	Q53020 r outer mem
22	166.5	6.1	2358	1	YEEJ_ECOLI	P76347 escherichia
23	166	6.1	550	1	FLIC_SHIFL	Q08860 shigella fl
24	165	6.0	1861	1	APU_THETU	P38536 t amylopull
25	164.5	6.0	918	1	YMJB_CAEEL	P34487 caenorhabdi
26	163.5	6.0	3178	1	YS89_CAEEL	Q09624 caenorhabdi
27	161.5	5.9	2334	1	WAPA_BACSU	Q07833 bacillus su
28	160	5.9	1005	1	Y456_CHLTR	O84462 chlamydia t
29	160	5.9	1770	1	PMPC_CHLTR	O84419 chlamydia t
30	159.5	5.8	948	1	HPI1_DEIRA	P56867 deinococcus
31	159	5.8	1286	1	AIDA_ECOLI	Q03155 escherichia
32	159	5.8	1694	1	IGA0_HAEIN	P44969 haemophilus
33	159	5.8	1702	1	IGA2_HAEIN	P45384 haemophilus
34	156.5	5.7	1007	1	Y741_CHLMU	Q9pjt6 chlamydia m
35	156	5.7	814	1	SLA1_BACAN	P49051 bacillus an
36	155.5	5.7	504	1	FLIC_SALBU	Q06969 salmonella
37	155.5	5.7	928	1	PM10_CHLPN	Q9rb65 chlamydia p
38	155	5.7	1310	1	VAC3_HELPY	Q48253 helicobacte
39	154.5	5.6	512	1	FLAA_HELFE	Q9xb38 helicobacte
40	154.5	5.6	575	1	FLA2_CAMJE	P22251 campylobact
41	154.5	5.6	928	1	HXA2_HAEIN	P45354 haemophilus
42	154	5.6	917	1	HXA3_HAEIN	P45355 haemophilus
43	153	5.6	582	1	FIBP_ADE02	P03275 human adeno
44	152.5	5.6	1723	1	PM20_CHLPN	Q9z812 chlamydia p
45	152	5.6	788	1	SP1_RAT	Q01714 rattus norv

ALIGNMENTS

RESULT 1
OMPA_RICRI
ID OMPA_RICRI STANDARD; PRT; 2249 AA.
AC P15921;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Outer membrane protein A precursor (190 kDa antigen) (Cell surface
DE antigen) (rOmpA) (rOmp A).
GN OMPA.
OS Rickettsia rickettsii.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Rickettsiaceae; Rickettsieae; Rickettsia.
OX NCBI_TaxID=783;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=R;
RX MEDLINE=90354033; PubMed=2117568;
RA Anderson B.E., McDonald G.A., Jones D.C., Regnery R.L.;
RT "A protective protein antigen of Rickettsia rickettsii has tandemly
RT repeated, near-identical sequences.";
RL Infect. Immun. 58:2760-2769(1990).
CC -!- FUNCTION: ELICITS PROTECTIVE IMMUNITY.
CC -!- SUBCELLULAR LOCATION: CELL WALL. THIS BACTERIUM IS COVERED BY A
CC S-LAYER WITH HEXAGONAL SYMMETRY.
CC -!- PTM: GLYCOSYLATED (PROBABLE).
CC -!- SIMILARITY: BELONGS TO THE RICKETTSIAE OMPA/OMPB FAMILY.
CC ---

CC ---
DR EMBL; M31227; AAA26380.1; -.
DR PIR; A41477; A41477.
DR InterPro; IPR006315; Autotransport.
DR InterPro; IPR005546; Autotransporter.
DR Pfam; PF03797; Autotransporter; 1.
DR TIGRFAMs; TIGR01414; autotrans_barl; 3.
KW Antigen; Repeat; Signal; Cell wall; S-layer; Glycoprotein.
FT SIGNAL 1 28 POTENTIAL.
FT CHAIN 29 2249 OUTER MEMBRANE PROTEIN A.
FT DOMAIN 212 1180 13 X APPROXIMATE TANDEM REPEATS.
FT REPEAT 212 286 A (TYPE I).
FT REPEAT 287 358 B (TYPE II).
FT REPEAT 359 430 C (TYPE II).
FT REPEAT 431 505 D (TYPE I).
FT REPEAT 506 577 E (TYPE II).
FT REPEAT 578 652 F (TYPE I).
FT REPEAT 653 724 G (TYPE II).
FT REPEAT 725 799 H (TYPE I).
FT REPEAT 800 874 I (TYPE I).
FT REPEAT 875 949 J (TYPE I).

```
FT   REPEAT      950   1021       K (TYPE II).
FT   REPEAT     1022   1093       L (TYPE II).
FT   REPEAT     1094   1165       M (TYPE II).
FT   REPEAT     1166   1180       TYPE I (INCOMPLETE).
SQ   SEQUENCE   2249 AA;  224333 MW;  A9D6646C089DF087 CRC64;

  Query Match             7.7%;  Score 210;  DB 1;  Length 2249;
  Best Local Similarity   23.8%;  Pred. No. 0.0069;
  Matches  139;  Conservative   68;  Mismatches  255;  Indels  122;  Gaps   27;

Qy          9 LESVQRSVVGSIQASMEGSVELETISLSMTNDSKEFVDPYIVVTLKAGDNLKIKQNTNEN 68
              |   | :|||||    :|::                    : ||| || :| :  |   |
Db        114 LNITQNTVVGSI--ITKGNL-------------------LPVTLNAGKSLTLNGN---N 148

Qy         69 TNASSFTYSLKKD-LTGLINVETEKLSFGANG---------KKVNIIS--DTKGLNFAKE 116
               |::  :    |  ||| |:       |||           |: :    |  |:   |
Db        149 AVAANHGFDAPADNYTGLGNIALG----GANAALIIQSAAPSKITLAGNIDGGGIITVKT 204

Qy        117 TAGTNGDTTVHLNGIGSTLTDMLLNTGATTNVTNDNV---TDDEKKRAASVKDVLNAGWN 173
              |  ||       ||:|     :| || |       |   |  :   ||||  : ||
Db        205 DAAINGT-------IGNTNALATVNVGAGTATLGGAVIKATTTKLTNAASVLTLTNANAV 257

Qy        174 IKGVKPGTTASDNVDFVRTYDTVEFLSADTKTT----TVNVESKD---NG---KKTEVKI 223
              : |    ||  |||  :     :  :: |   |    |::| :      |   | |  |:
Db        258 LTGAIDNTTGGDNVGVLNLNGALSQVTGDIGNTNSLATISVGAGTATLGGAVIKATTTKL 317

Qy        224 GAKTSVIKEKDGKLVTG----KGKGENGSSTDEGEGLVTAK----EVIDAVNKAGWRMKT 275
              | :|  :  :|||     |   ||  |  |  ||        :  ||     ::
Db        318 TDAASAVKFTNPVVVTGAIDNTGNANNGIVTFTGNSTVTGNVGNTNALATVNVGAGLLQ- 376

Qy        276 TTANGQTGQADKFETVTSGTKVTFASGNGTTATVSKDDQGNIT---VKYDVN---VGDAL 329
              |  :|:       : : |||   |    | : |: ||     |  :  |    ||
Db        377 --VQGGVVKANTINLTDNASAVTFT--NPVVVTGAIDNTGNANNGIVTFTGNSTVTGDIG 432

Qy        330 NVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINA--GNNIEITRNGKNID 387
              | | |        :  : |:  |  :  :: :   :  | | ||     |: |  | |:
Db        433 NTNALATVNVGAGTATLGGAVIKATTTKLTNAASVLTLT-NANAVLTGAIDNTTGGDNVG 491

Qy        388 IAT-----------SMTPQFSSVSLGAGADAPTLSVDDEGALNVGSKDANKPVRITNVA 435
              :                |   :::|:|||    ||       |      ||   |: ||
Db        492 VLNLNGALSQVTGNIGNTNSLATISVGAG--TATLGGAVIKATTTKLTDAASAVKFTN-- 547

Qy        436 PGVKEGDVTNVAQLKGVAQNLNNRIDNVNGNARA----GIAQAIATAGLVQAYLPGKSMM 491
              | |  | : |          | || |    ||:      |   ::||  :      |
Db        548 PVVVTGAIDNTG-------NANNGIVTFTGNSTVTGDIGNTNSLATISV------GAGTA 594

Qy        492 AIGGG------TYLGEAGYAIGYSSISAGGNWIIKGTASGNSRG 529
              :||       | |  |   :  :: :|     |  |  |:: |
Db        595 TLGGAVIKATTTKLTNAASVLTLTNANAVLTGAIDNTTGGDNVG 638


RESULT 2
OMPB_RICCN
ID   OMPB_RICCN     STANDARD;      PRT;  1655 AA.
AC   Q9KKA3; Q9KK98; Q9XC45;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Outer membrane protein B precursor (168 kDa surface-layer protein)
DE   (Surface protein antigen) (Cell surface antigen 5) (Sca5) (rOmpB)
DE   (rOmp B) [Contains: 120 kDa surface-exposed protein (Surface protein
DE   antigen) (120 kDa outer membrane protein ompB); 32 kDa beta peptide].
GN   OMPB OR RC1085.
OS   Rickettsia conorii.
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC   Rickettsiaceae; Rickettsieae; Rickettsia.
OX   NCBI_TaxID=781;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Malish 7;
RX   MEDLINE=21442074; PubMed=11557893;
RA   Ogata H., Audic S., Renesto-Audiffren P., Fournier P.-E., Barbe V.,
RA   Samson D., Roux V., Cossart P., Weissenbach J., Claverie J.-M.,
RA   Raoult D.;
RT   "Mechanisms of evolution in Rickettsia conorii and R. prowazekii.";
RL   Science 293:2093-2098(2001).
RN   [2]
RP   SEQUENCE OF 33-1649 FROM N.A.
RC   STRAIN=Indian tick typhus, and Malish 7;
RX   MEDLINE=20393643; PubMed=10939649;
RA   Roux V., Raoult D.;
RT   "Phylogenetic analysis of members of the genus Rickettsia using the
RT   gene coding the outer-membrane protein rOmpB (ompB).";
RL   Int. J. Syst. Evol. Microbiol. 50:1449-1455(2000).
RN   [3]
RP   SEQUENCE OF 353-1655 FROM N.A.
RC   STRAIN=Malish 7;
RA   Stenos J., Walker D.;
RT   "The rickettsial outer membrane protein A and B genes of Rickettsia
RT   australis, the most divergent rickettsia of the spotted fever group.";
RL   Submitted (MAY-1999) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: THE 120 KDA SURFACE-EXPOSED PROTEIN IS A MAJOR
CC       STRUCTURAL PROTEIN WHICH MAY PLAY A ROLE AS A RICKETTSIAL
CC       VIRULENCE FACTOR AND/OR IMMUNOGEN DURING INFECTION (BY
CC       SIMILARITY).
CC   -!- FUNCTION: THE 32 KDA BETA PEPTIDE MAY SERVE AS A MEMBRANE ANCHOR
CC       (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: CELL WALL. THIS BACTERIUM IS COVERED BY A S-
CC       LAYER WITH HEXAGONAL SYMMETRY (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE RICKETTSIAE OMPA/OMPB FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE008659; AAL03623.1; -.
DR   EMBL; AF123721; AAF34124.1; -.
DR   EMBL; AF123726; AAF34129.1; -.
DR   EMBL; AF149110; AAD39533.1; -.
DR   PIR; E97835; E97835.
DR   InterPro; IPR006315; Autotransport.
DR   InterPro; IPR005546; Autotransporter.
DR   Pfam; PF03797; Autotransporter; 1.
DR   TIGRFAMs; TIGR01414; autotrans_barl; 2.
KW   Antigen; S-layer; Cell wall; Complete proteome.
FT   CHAIN         1   1334       120 kDa SURFACE-EXPOSED PROTEIN.
FT   CHAIN      1335   1655       32 kDa BETA PEPTIDE.
FT   VARIANT      61     61       P -> A (IN STRAIN INDIAN TICK TYPHUS).
FT   VARIANT      75     75       G -> S (IN STRAIN INDIAN TICK TYPHUS).
FT   VARIANT      78     78       K -> N (IN STRAIN INDIAN TICK TYPHUS).
FT   VARIANT     251    251       V -> A (IN STRAIN INDIAN TICK TYPHUS).
FT   VARIANT     413    413       N -> D (IN STRAIN INDIAN TICK TYPHUS).
FT   VARIANT     959    959       I -> V (IN STRAIN INDIAN TICK TYPHUS).
FT   VARIANT     988    988       A -> T (IN STRAIN INDIAN TICK TYPHUS).
FT   VARIANT    1139   1139       R -> L (IN STRAIN INDIAN TICK TYPHUS).
FT   CONFLICT    353    354       KD -> GH (IN REF. 3).
FT   CONFLICT    776    776       F -> S (IN REF. 3).
FT   CONFLICT   1159   1159       E -> D (IN REF. 3).
FT   CONFLICT   1177   1177       G -> S (IN REF. 3).
FT   CONFLICT   1492   1492       H -> R (IN REF. 3).
SQ   SEQUENCE   1655 AA;  168342 MW;  E49E19377D5FCE37 CRC64;

  Query Match             7.6%;  Score 206.5;  DB 1;  Length 1655;
  Best Local Similarity   23.0%;  Pred. No. 0.0072;
  Matches  134;  Conservative   73;  Mismatches  251;  Indels  125;  Gaps   28;

Qy         15 SVVGSIQASMEGSVELETISLSMTNDSKEFVDPYIVVTLKAGDNLKIKQNTNENTNASSF 74
                :| ||:  :::| :|  |::::              :   || |  :    |    |: |
Db        287 NVTGSLGGNLKGIIEFNTVAVNGQ-----------LKANAGANAAVIGTNNGAGRAAGF 334

Qy         75 TYSLKKDLTGLINVETEKLSFGANGKKVNIISDTKGLNFAKETAGTNGDTTV-HLNGIGS 133
               |:                  |||   |       :   ::|   |    |:  :|:
```

```
Db        335 VVSVD-----------------NGKVATIDGQVYAKDMVIQSANAVGQVNFRHIVDVGT 376

Qy        134 TLTDMLLNTGATTNVT-NDNV-TDDEKKRAASV----KDVLNAGWNIKGVKPGTTASDNV 187
Db        377 DGTTAFKTAASKVAITQNSNFGTTDFGNLAAQIIVPNTMTLNGNFTGDASNPGNTAG--- 433

Qy        188 DFVRTYD---TVEFLSADTKTTTVN----VESKDNG-------KKTEVKIGAKTSVIKEK 233
Db        434 --VITFDANGTLASASADANVAVTNNITAIEASGAGVVQLSGTHAAELRLGNAGSVFKLA 491

Qy        234 DGKLVTGK--------GKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQA 285
Db        492 DGTVINGKVNQTALVGGALAAGTITLDGSATITG----DIGNAGG----AAALQGITLAN 543

Qy        286 DKFETVT---------SGTKVTFASGNGTTATVSKDDQGNITVKYDV-------NVGDAL 329
Db        544 DATKTLTLGGANIIGANGGTINFQANGGTIKLTS--TQNNIVVDFDLAIATDQTGVVDAS 601

Qy        330 NVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKM-DETVNIN---AGNN--IEITRNG 383
Db        602 SLTNAQTLTINGKIGTVGANNKTLGQFNIGSSKTVLSDGDVAINELVIGNNGAVQFAHNT 661

Qy        384 KNIDIATSMTPQ----FSSV-----SLGAGADAPTLSVDDEGALNVGSKDANKPVRITNV 434
Db        662 YLITRTTNAAGQGKIIFNPVVNNNTTLATGTNLGS-ATNPLAEINFGSKGAANVDTVLNV 720

Qy        435 APGVKEGDVTNVAQLKGVAQNLNNRIDNVNG-NARAGIAQAIATAGLVQ-------AYLP 486
Db        721 GKGVNL-YATNITTTDA---NVGSFIFNAGGTNIVSG------TVGGQQGNKFNTVALDN 770

Qy        487 GKSMMAIGGGTYLGEAGYAIGYSSISAGGNWIIKGTASGNSRG 529
Db        771 GTTVKFLGNATFNGNTTIAAN-STLQIGGNYTADFVASADGTG 812
```

RESULT 3
OMPB_RICRI

```
ID   OMPB_RICRI     STANDARD;      PRT;  1654 AA.
AC   Q53047;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Outer membrane protein B precursor (168 kDa surface-layer protein)
DE   (Surface protein antigen) (Cell surface antigen 5) (Sca5) (rOmpB)
DE   (rOmp B) [Contains: 120 kDa surface-exposed protein (Surface protein
DE   antigen) (120 kDa outer membrane protein ompB); 32 kDa beta peptide].
GN   OMPB.
OS   Rickettsia rickettsii.
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC   Rickettsiaceae; Rickettsieae; Rickettsia.
OX   NCBI_TaxID=783;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=R;
RX   MEDLINE=92167802; PubMed=1724278;
RA   Gilmore R.D. Jr., Cieplak W. Jr., Policastro P.F., Hackstadt T.;
RT   "The 120 kilodalton outer membrane protein (rOmp B) of Rickettsia
RT   rickettsii is encoded by an unusually long open reading frame:
RT   evidence for protein processing from a large precursor.";
RL   Mol. Microbiol. 5:2361-2370(1991).
RN   [2]
RP   SEQUENCE OF 279-1654 FROM N.A.
RC   STRAIN=R;
RX   MEDLINE=90136087; PubMed=2515418;
RA   Gilmore R.D. Jr., Joste N., McDonald G.A.;
RT   "Cloning, expression and sequence analysis of the gene encoding the
RT   120 kD surface-exposed protein of Rickettsia rickettsii.";
RL   Mol. Microbiol. 3:1579-1586(1989).
CC   -!- FUNCTION: THE 120 kDa SURFACE-EXPOSED PROTEIN IS A MAJOR
CC       STRUCTURAL PROTEIN WHICH MAY PLAY A ROLE AS A RICKETTSIAL
CC       VIRULENCE FACTOR AND/OR IMMUNOGEN DURING INFECTION.
CC   -!- SUBCELLULAR LOCATION: CELL WALL. THIS BACTERIUM IS COVERED BY A S-
CC       LAYER WITH HEXAGONAL SYMMETRY.
CC   -!- SIMILARITY: BELONGS TO THE RICKETTSIAE OMPA/OMPB FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; X16353; CAA34403.1; -.
DR   PIR; S18227; S18227.
DR   InterPro; IPR006315; Autotransport.
DR   InterPro; IPR005546; Autotransporter.
DR   Pfam; PF03797; Autotransporter; 1.
DR   TIGRFAMs; TIGR01414; autotrans_barl; 2.
KW   Antigen; S-layer; Cell wall.
FT   CHAIN         1   1333       120 kDa SURFACE-EXPOSED PROTEIN.
FT   CHAIN      1334   1654       32 kDa BETA PEPTIDE.
FT   DOMAIN     1181   1188       POLY-THR.
SQ   SEQUENCE   1654 AA;  168184 MW;  D7AB70FB7087F618 CRC64;

  Query Match             7.3%;  Score 201;  DB 1;  Length 1654;
  Best Local Similarity   22.8%;  Pred. No. 0.013;
  Matches 136;  Conservative   79;  Mismatches  228;  Indels  154;  Gaps   32;

Qy         15 SVVGSIQASMEGSVELETISLSMTNDSKEFVDPYIVVTLKAGDNLKIKQNTNENTNASSF 74
Db        287 NITGSLGGNLKGVIEFNTVAV----DGQ--------LTANAGAANAVIGTNNGAGRAAGF 334

Qy         75 TYSLKKDLTGLINVETEKLSFGANGKKVNIISDTKGLNFAKETAGTNGDTTV-HLNGIGS 133
Db        335 VVSVD-----------------NGKVATIDGQVYAKDMVIQSANATGQVNFRHIVDVGA 376

Qy        134 TLTDMLLNTGATTNVTNDN--VTDDEKKRAASVK--DVLNAGWNIKG--VKPGTTASDNV 187
Db        377 DGTTAFKTAASKVTITQDSNFGNTDFGNLAAQIKVPNAITLTGNFTGDASNPGNTAG--- 433

Qy        188 DFVRTYD---TVEFLSADTKTTTVN----VESKDNG-------KKTEVKIGAKTSVIKEK 233
Db        434 --VITFDANGTLESASADANVAVTNNITAIEASGAGVVQLSGTHAAELRLGNAGSIFKLA 491

Qy        234 DGKLVTGK--------GKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQA 285
Db        492 DGTVINGKVNQTALVGGALAAGTITLDGSATITGD--IGNAGGAAALQRITLAN------ 543

Qy        286 DKFETVTSGTKVTFASGNGT-----------TATVSKDDQGNITVKYDV-------NVG 326
Db        544 DAKKTLTLGGANIIGAGGGTIDLQANGGTIKLTST-----QNNIVVDFDLAIATDQTGVV 598

Qy        327 DALNVNQLQNSGWN-------LDSKAVA----GSSGKVIS-GNVSPSKGKMDETVNINAG 374
Db        599 DASSLTNAQTLTINGKIGTIGANNKTLGQFNIGSSKTVLSNGNVA-----INELVIGNDG 653

Qy        375 ------NNIEITR------NGKNI--DIATSMTPQFSSVSLGAGADAPTLSVDDEGALNV 420
Db        654 AVQFAHDTYLITRTTNAAGQGKIIFNPVVNNGTTLAAGTNLGS-------ATNPLAEINF 706

Qy        421 GSKDANKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVNG-NARAGIAQAIATAG 479
Db        707 GSKGVNVDT-VLNVGEGVNL-YATNITTTDA---NVGSFVFNAGGTNIVSG------TVG 755

Qy        480 LVQ-------AYLPGKSMMAIGGGTYLGEAGYAIGYSSISAGGNWIIKGTASGNSRG 529
Db        756 GQQGNKFNTVALENGTTVKFLGNATFNGNTTIAAN-STLQIGGNYTADCVASADGTG 811
```

RESULT 4
YDBA_ECOLI

```
ID   YDBA_ECOLI     STANDARD;      PRT;  2003 AA.
AC   P33666; P76087; P76088; P76856; P76857; P76859;
DT   01-FEB-1994 (Rel. 28, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
```

```
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Hypothetical protein ydbA.
GN   YDBA OR B1401/B1405.
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12 / MG1655;
RX   MEDLINE=97426617; PubMed=9278503;
RA   Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA   Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA   Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA   Mau B., Shao Y.;
RT   "The complete genome sequence of Escherichia coli K-12.";
RL   Science 277:1453-1474(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12;
RX   MEDLINE=97251357; PubMed=9097039;
RA   Aiba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K., Itoh T.,
RA   Kasai H., Kashimoto K., Kimura S., Kitakawa M., Kitagawa M.,
RA   Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,
RA   Nakade S., Nakamura Y., Nashimoto H., Nishio Y., Oshima T., Saito N.,
RA   Sampei G., Seki Y., Sivasundaram S., Tagami H., Takeda J.,
RA   Takemoto K., Takeuchi Y., Wada C., Yamamoto Y., Horiuchi T.;
RT   "A 570-kb DNA sequence of the Escherichia coli K-12 genome
RT   corresponding to the 28.0-40.1 min region on the linkage map.";
RL   DNA Res. 3:363-377(1996).
RN   [3]
RP   SEQUENCE OF 464-2003 FROM N.A.
RC   STRAIN=K12;
RX   MEDLINE=92190338; PubMed=1665988;
RA   Moszer I., Glaser P., Danchin A.;
RT   "Multiple IS insertion sequences near the replication terminus in
RT   Escherichia coli K-12.";
RL   Biochimie 73:1361-1374(1991).
CC   -!- SIMILARITY: TO S.TYPHIMURIUM ORF NEAR CYSG (AC P25928).
CC   -!- CAUTION: THIS IS A CONCEPTUAL TRANSLATION; THE GENE CODING FOR
CC       THIS PROTEIN IS INTERRUPTED BY A HYBRID IS2D/IS30C ELEMENT
CC       BETWEEN AMINO ACIDS 839 AND 840.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AE000237; AAC74483.1; ALT_SEQ.
DR   EMBL; AE000237; AAC74487.1; ALT_SEQ.
DR   EMBL; D90778; BAA15009.1; ALT_SEQ.
DR   EMBL; D90778; BAA18880.1; ALT_SEQ.
DR   EMBL; D90779; BAA18881.1; ALT_SEQ.
DR   EMBL; X62680; -; NOT_ANNOTATED_CDS.
DR   EcoGene; EG11307; ydbA.
KW   Hypothetical protein; Complete proteome.
FT   CONFLICT    489    489       I -> V (IN REF. 2).
FT   CONFLICT    495    495       I -> V (IN REF. 2).
SQ   SEQUENCE   2003 AA;  205949 MW;  B83A12C8B53220EE CRC64;

  Query Match             7.2%;  Score 198;  DB 1;  Length 2003;
  Best Local Similarity   23.1%;  Pred. No. 0.023;
  Matches  158;  Conservative   88;  Mismatches  239;  Indels  198;  Gaps   39;

Qy          4 DEEEELESVQRSVVGSIQASMEGSVELE-------TISLSMTNDSKEFV----------- 45
              |: |:  ::: ||  :   : :|:: |:       ||:|   :::   |
Db        149 DKTEKTLTIRDSVF-TYTENADGTISLQDSNGRKATINLWQIDEANNTVALEGVSADGAT 207

Qy         46 ------DPYIVVTLKAGDNLKIKQNTNENTNASSFTYSLKKDLTGL-IN------VETEK 92
                    :  :|:|   |||. :  | |  |          || ||  ||      ::

Db        208 KWQYNHNGELVIT---GDNATV--NNNGKTTVDG------KDSTGTEINGNNGKVIQDGD 256

Qy         93 LSFGANGKKVNIISDT-----KG--------------------LNFAKETAGTNGDTTVH 127
              |     |  ::|  |:     ||                    :|   |:  ||| |
Db        257 LDVSGGGHGIDITGDSATVDNKGTMTVTDPESMGIQIDGDKAIVNNEGESTITNGGTGTQ 316

Qy        128 LNG-------IGSTLTDMLLNTGATTNVTNDNVTDD--------------------EKK 159
              :||        | |  |   :||   |  |  |  |                      : |
Db        317 INGDDATANNNGKTTVDGKDSTGTEINGNNGKVIQDGDLDVSGGGHGIDITGDSATVDNK 376

Qy        160 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKK- 218
                :| |  : |  : |       ::  :    |   :: |  |   | ::  :||
Db        377 GTMTVTDPESIGIQVDG-DQAVVNNEGESAITNGGTGTQINGDDATANNNGKTTVDGKDS 435

Qy        219 --TEVKIGAKTSVIKEKDGKL-VTGKGKG---ENGSSTDEGEGLVTAKE------VIDA- 265
                ||:  |    ||  :|| | |:| | |      |:| : :| :|  :       ||
Db        436 TGTEI-AGNNGKVI--QDGDLDVSGGGHGIDITGDSATVDNKGTMTVTDPESIGIQIDGD 492

Qy        266 ---VNKAGWRMKTTTANGQTGQADKFETVTSGTKVTFASGNGTTATVSKDD-----QGNI 317
                ||  |   ::|  || ||      |  :|   | |: :| |    ||       |||
Db        493 QAIVNNEG---ESTITNGGTG------TQINGNDAT-ANNSGKTTVDGKDSTGTKIAGNI 542

Qy        318 TVKYDVNVGDALNV----NQLQNSGWN--LDSK-----AVAGSSGKVISGNVSPSKGKMD 366
               :   ||:  :| |    : ::| | |  :::|     :  || | :|:|  :      |
Db        543 GI---VNLDGSLTVTGGAHGVENIGDNGTVNNKGDIVVSDTGSIGVLINGEGATVSNTGD 599

Qy        367 ETVNINAGNNIEITRNGKNIDIATSM-TPQFSS-VSLGAGADAPTLSVDD-------EGA 417
                |: |       || |   : :| ||    ||: | |    :: ||:  |
Db        600 VNVS-NEATGFSITTNSGKVSLAGSMQVGDFSTGVDLNGNNNSVTLAAKDLKVVGQKATG 658

Qy        418 LNVGSKDANKPVRITNVAPGVKEGDVTNVAQL---KGVAQNLNNRIDNVNGNARAGI--- 471
              :|| | |||  | ||       |:     | |:       |  |:    :||  : :  :
Db        659 INV-SGDAN-TVNITGNVLVDKDKTADNAAEYFFDPSVGINVYGSDNNVTLDGKLTVVSD 716

Qy        472 ------------AQAIATAGLVQAYLPGKSMMAIGGGTYLGE--------------AGYA 505
                         |  |:|||     | ::   ||   :||                ||:
Db        717 SEVTSRQSNLFDGSAEKTSGLV-VIGDGNTVNMNGGLELIGEKNALADGSQVTSLRTGYS 775

Qy        506 I-------GYSSISAGGNWIIKG 521
                    | ||:   |:  | |
Db        776 YTSVIVVSGESSVYLNGDTTISG 798


RESULT 5
HLYA_PROMI
ID   HLYA_PROMI     STANDARD;      PRT;  1577 AA.
AC   P16466;
DT   01-AUG-1990 (Rel. 15, Created)
DT   01-AUG-1990 (Rel. 15, Last sequence update)
DT   01-NOV-1990 (Rel. 16, Last annotation update)
DE   Hemolysin precursor.
GN   HPMA.
OS   Proteus mirabilis.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Proteus.
OX   NCBI_TaxID=584;
RN   [1]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 30-43.
RC   STRAIN=Isolate 477-12;
RX   MEDLINE=90170827; PubMed=2407716;
RA   Uphoff T.S., Welch R.A.;
RT   "Nucleotide sequencing of the Proteus mirabilis calcium-independent
RT   hemolysin genes (hpmA and hpmB) reveals sequence similarity with the
RT   Serratia marcescens hemolysin genes (shlA and shlB).";
RL   J. Bacteriol. 172:1206-1216(1990).
CC   -!- FUNCTION: BACTERIAL HEMOLYSINS ARE EXOTOXINS THAT ATTACK BLOOD
CC       CELL MEMBRANES AND CAUSE CELL RUPTURE BY MECHANISMS NOT CLEARLY
CC       DEFINED.
CC   -!- FUNCTION: CELL-BOUND HEMOLYSIN, WHICH RELEASES HEME-IRON FROM
CC       ERYTHROCYTES BY INTERACTION WITH THE ERYTHROCYTE MEMBRANE. HPMA
CC       REQUIRES HPMB FUNCTION.
CC   -!- SUBCELLULAR LOCATION: Outer membrane.
CC   -!- MISCELLANEOUS: THE CONSERVED AMPHIPATHIC DOMAINS IN SHLA AND HPMA
```

```
CC       MAY BE RESPONSIBLE FOR PORE FORMATION.
CC    -!- SIMILARITY: TO S.MARCESCENS HEMOLYSIN (SHLA).
CC    ---------------------------------------------------------------------
CC    This SWISS-PROT entry is copyright. It is produced through a collaboration
CC    between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC    the European Bioinformatics Institute. There are no restrictions on its
CC    use by non-profit institutions as long as its content is in no way
CC    modified and this statement is not removed. Usage by and for commercial
CC    entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC    or send an email to license@isb-sib.ch).
CC    ---------------------------------------------------------------------
DR    EMBL; M30186; AAA25657.1; -.
DR    PIR; A35140; A35140.
KW    Hemolysis; Toxin; Outer membrane; Signal.
FT    SIGNAL        1     29
FT    CHAIN        30   1577       HEMOLYSIN.
SQ    SEQUENCE   1577 AA;  165869 MW;  175975E0C924B2D9 CRC64;

  Query Match             7.2%;  Score 197;  DB 1;  Length 1577;
  Best Local Similarity   22.3%;  Pred. No. 0.02;
  Matches  150;  Conservative   94;  Mismatches  190;  Indels  240;  Gaps   39;

Qy         15 SVVGSI-----------------QASMEGSVELETISLSMTNDSKEFVDPYIVVTLKAG 56
              :|:||:                  |     :: |  ||::|  :||  |        ||
Db        644 NVIGSLIKSADKLGIHSLGDINVKSAQQVTKIDDEKTSLAITGHAKEVEDK----QYSAG 699

Qy         57 DNLKIKQNTNENTNASSFTYSLKKDLTGLINVETEKLSFGANGKKVNIISDTKGLNFA-- 114
                 |  ||:|| :                : |||: :   :|  |:: :: | : ||
Db        700 --FHITHTTNKNT----------------STETEQANSTISGANVDLQAN-KDVTFAGS 739

Qy        115 --KETAGT--------------------NGDTTVH--------LNGIGS----------T 134
               | |||                     | |||:         :: :||          |
Db        740 DLKTTAGNASITGDNVAFVSTENKKQTDNTDTTISGGFSYTGGVDKVGSKADFQYDKQHT 799

Qy        135 LTDMLLNTGATTNVTND-NVTDDEKKRAASVKDVLNAGWNIKGVKPGTTASDNVD--FVR 191
              |::  | |: | |  |  :| :        ||:|: |           || :|:  :
Db        800 QTEVTKNRGSQTEVAGDLTITAN--------KDLLHEG-----------ASHHVEGRYQE 840

Qy        192 TYDTVEFL----SADTKTTTVNV-----------------ESKDNGKKTEVKIGAKTSV 229
              : : :: |    |  :|| ::||                 :: ::|  |  | |  | :
Db        841 SGENIQHLAVNDSETSKTDSLNVGIDVGVNLDYSGVTKPVKKAIEDGVNT-TKPGNNTDL 899

Qy        230 IKE---KD---------------GKLVTGKGKGENGSSTDE---GEGLVTAKEVIDAVN 267
              |:   :|               |  |  || |   | ||       :   |  ||: |
Db        900 TKKVTARDAIANLANLSNLETPNVGVEVGIKGGGSQQSQTDSQAVSTSINAGKIDIDSNN 959

Qy        268 K---AGWRMKTT------TANGQTGQA--DKFETVTSGTK----VTFASGNGTTATVSKD 312
              |    |   ::|      |||  | :|  || :|    ||    :  ::  |:  ||:
Db        960 KLHDQGTHYQSTQEGISLTANTHTSEATLDKHQTTFHETKGGGQIGVSTKTGSDITVAIK 1019

Qy        313 DQGNITVKYDVNVGDALNVNQLQNSGW--NLDSKAVAGSSGKVISGNVSPSKGKMDETVN 370
              :|  |       :||   : : | :  | |     | :           |||   ||
Db       1020 GEGQTT-------DNALMETKAKGSQFTSNGDISINVGENAHYEGAQFDAQKGK---TV- 1068

Qy        371 INAGNNIEITR-----------NGK-NIDIATSMTPQFSSVSLGAGADAPTL-SVDDEG 416
              |||| :: : :           ||  |: : |  ||:  |   | | :| |     ::
Db       1069 INAGGDLTLAQATDTHSESQSNVNGSANLKVGT--TPE--SKDYGGGFNAGTTHHSKEQT 1124

Qy        417 ALNVGSKDANKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRID-------------NV 463
                ||:   :: : :        :|  |::  : :| |  |::|              |:
Db       1125 TAKVGTITGSQGIELNAGHNLTLQG--THLSSEQDIALNATNKVDLQSASSEHTEKGNNL 1182

Qy        464 NGNARAGIAQAIATAGLVQAYLPGKSM----MAIGGGTYLGEAGYAIG---YSSISAGGN 516
              :|  :||              || |     :: |   || | :|||     |:|  |
Db       1183 SGGVQAGF--------------GKKMTDDASSVNG---LGSAQFAIGKQDEKSVSREGG 1224

Qy        517 WIIKG---TASGNS 527
              |      | :|||
Db       1225 TINNSGNLTINGNS 1238


RESULT 6
OMPA_RICCN
ID   OMPA_RICCN     STANDARD;      PRT;  2021 AA.
AC   Q52657; P95591; P95592; P95593; P95594; Q52667; Q52668; Q52669;
AC   Q52670; Q52674;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Outer membrane protein A precursor (190 kDa antigen) (Cell surface
DE   antigen) (rOmpA) (rOmp A).
GN   OMPA OR RC1273.
OS   Rickettsia conorii.
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC   Rickettsiaceae; Rickettsieae; Rickettsia.
OX   NCBI_TaxID=781;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Malish 7;
RX   MEDLINE=94171067; PubMed=8125327;
RA   Crocquet-Valdes P.A., Weiss K., Walker D.H.;
RT   "Sequence analysis of the 190-kDa antigen-encoding gene of Rickettsia
RT   conorii (Malish 7 strain).";
RL   Gene 140:115-119(1994).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Malish 7;
RX   MEDLINE=21442074; PubMed=11557893;
RA   Ogata H., Audic S., Renesto-Audiffren P., Fournier P.-E., Barbe V.,
RA   Samson D., Roux V., Cossart P., Weissenbach J., Claverie J.-M.,
RA   Raoult D.;
RT   "Mechanisms of evolution in Rickettsia conorii and R. prowazekii.";
RL   Science 293:2093-2098(2001).
RN   [3]
RP   SEQUENCE OF 8-204 FROM N.A.
RC   STRAIN=Indian tick typhus, M1, Malish 7, and Moroccan;
RX   MEDLINE=97015921; PubMed=8862558;
RA   Roux V., Fournier P.E., Raoult D.;
RT   "Differentiation of spotted fever group rickettsiae by sequencing and
RT   analysis of restriction fragment length polymorphism of PCR-amplified
RT   DNA of the gene encoding the protein rOmpA.";
RL   J. Clin. Microbiol. 34:2058-2065(1996).
RN   [4]
RP   SEQUENCE OF 953-2012 FROM N.A.
RC   STRAIN=Indian tick typhus, M1, Malish 7, and Moroccan;
RA   Raoult D., Fournier P.E., Roux V.;
RT   "Phylogenetic analysis of spotted fever group rickettsiae by study
RT   of the outer surface protein rOmpA.";
RL   Submitted (DEC-1996) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: ELICITS PROTECTIVE IMMUNITY (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: CELL WALL. THIS BACTERIUM IS COVERED BY A
CC       S-LAYER WITH HEXAGONAL SYMMETRY.
CC   -!- PTM: GLYCOSYLATED (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE RICKETTSIAE OMPA/OMPB FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U01028; AAA17405.1; -.
DR   EMBL; AE008674; AAL03811.1; -.
DR   EMBL; U43794; AAB49549.1; -.
DR   EMBL; U43798; AAB49550.1; -.
DR   EMBL; U43806; AAB49551.1; -.
DR   EMBL; U45244; AAB49566.1; -.
DR   EMBL; U46918; AAA86663.1; -.
DR   EMBL; U83440; AAC35176.1; -.
DR   EMBL; U83443; AAC35179.1; -.
DR   EMBL; U83448; AAC35184.1; -.
DR   EMBL; U83453; AAC35189.1; -.
DR   InterPro; IPR006315; Autotransport.
DR   Pfam; PF03797; Autotransporter; 1.
```

```
DR   TIGRFAMs; TIGR01414; autotrans_barl; 1.
KW   Antigen; Repeat; Signal; Cell wall; S-layer; Glycoprotein;
KW   Complete proteome.
FT   SIGNAL        1     38       POTENTIAL.
FT   CHAIN        39   2021       OUTER MEMBRANE PROTEIN A.
FT   DOMAIN      238    946       THR-RICH.
FT   DOMAIN     1424   1528       THR-RICH.
FT   VARIANT      60     60       N -> NN (IN STRAIN INDIAN TICK TYPHUS).
FT   VARIANT      76     76       R -> H (IN STRAIN INDIAN TICK TYPHUS).
FT   VARIANT      86    137       MISSING (IN STRAIN M1).
FT   VARIANT     126    133       MISSING (IN STRAIN MOROCCAN).
FT   VARIANT     953    954       VT -> II (IN STRAIN INDIAN TICK TYPHUS).
FT   VARIANT    1245   1245       D -> A (IN STRAINS INDIAN TICK TYPHUS, M1
FT                                AND MOROCCAN).
FT   VARIANT    1308   1308       N -> H (IN STRAIN MOROCCAN)..
FT   VARIANT    1877   1877       M -> I (IN STRAIN INDIAN TICK TYPHUS).
FT   CONFLICT     10     10       Q -> K (IN REF. 1).
FT   CONFLICT     92     92       I -> V (IN REF. 1).
FT   CONFLICT    126    126       V -> I (IN REF. 1).
FT   CONFLICT    137    137       T -> N (IN REF. 1).
FT   CONFLICT    157    157       G -> D (IN REF. 1).
FT   CONFLICT    368    369       IS -> VN (IN REF. 1).
FT   CONFLICT    374    388       KATLGGAIIKATTTK -> LLQVQGGVVKANTIN (IN
FT                                REF. 1).
FT   CONFLICT    640    640       N -> D (IN REF. 1).
FT   CONFLICT    669    669       V -> I (IN REF. 1).
FT   CONFLICT    793    793       N -> D (IN REF. 1).
FT   CONFLICT    803    804       VN -> IS (IN REF. 1).
FT   CONFLICT    809    823       LLRVQGGVVKSNTIN -> KATLGGAIIKATTTK (IN
FT                                REF. 1).
FT   CONFLICT    898    898       D -> Y (IN REF. 1).
FT   CONFLICT    908    908       P -> N (IN REF. 1).
FT   CONFLICT    985    985       N -> K (IN REF. 1).
FT   CONFLICT   1009   1009       L -> S (IN REF. 1).
FT   CONFLICT   1013   1013       Y -> S (IN REF. 1).
FT   CONFLICT   1182   1182       K -> Q (IN REF. 1).
FT   CONFLICT   1314   1314       N -> Y (IN REF. 4).
FT   CONFLICT   1451   1451       H -> N (IN REF. 1).
FT   CONFLICT   1624   1624       G -> D (IN REF. 1).
FT   CONFLICT   1628   1628       E -> G (IN REF. 1).
FT   CONFLICT   1872   1872       A -> V (IN REF. 1).
FT   CONFLICT   1875   1875       T -> P (IN REF. 1).
FT   CONFLICT   1878   1879       MS -> LP (IN REF. 1).
FT   CONFLICT   1936   1936       E -> A (IN REF. 1).
FT   CONFLICT   1965   1970       MTAPLP -> ITPPLS (IN REF. 1).
FT   CONFLICT   1997   1997       G -> R (IN REF. 1).
SQ   SEQUENCE   2021 AA;  203328 MW;  327FC42D7CB24668 CRC64;

  Query Match             6.9%;  Score 189.5;  DB 1;  Length 2021;
  Best Local Similarity   21.7%;  Pred. No. 0.061;
  Matches  155;  Conservative   81;  Mismatches  248;  Indels  231;  Gaps   34;

Qy         17 VGSIQASMEGS-VELETISLSMTNDSKEFVDPYIV------------VTLKAGDNLKIK 62
              ||: :|:: |: ::  |  |:    :  | :| :|            :    ||:
Db        514 VGAGKATLGGAIIKATTTKLTDNASAVTFTNPVVVTGAIDNTGNANNGIVTFTGDSTVTG 573

Qy         63 QNTNENTNASSFTYSLKKDLTG-LINVETEKLSFGANGKKVNIISDTKGLNFAKETAGTN 121
                | |  |:    : |  | | :|   | ||:   |   |   :          |   |
Db        574 NIGNTNALATISVGAGKATLGGAIIKATTTKLT--DNASAVTFTNPVVVTGAIDNTGNAN 631

Qy        122 -------GDTTVHLNGIGSTLTDMLLNTGA------------TTNVTNDNVTDDEKKRA 161
                   |::||  | ||:|     :| ||              || :||          |
Db        632 NGIVTFTGNSTVTGN-IGNTNALATVNVGAGIATLEGAVIKATTTKLTN----------A 680

Qy        162 ASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTT----TVNV---ESKD 214
              |||  : |    : |    ||  |||  :     :  :: :   |    |::|   ::
Db        681 ASVLTLTNVNAVLTGAIDNTTGVDNVGVLNLNGALSQVTGNIGNTNALATISVGAGKATL 740

Qy        215 NG---KKTEVKIGAKTSVIKEKDGKLVTG----KGKGENGSSTDEGEGLVTAK----EVI 263
                |  | |  |:    | :   :  :|||     |   || :|  |:  ||        :
Db        741 GGAVIKATTTKLTDNASAVTFTNPVVVTGAIDNTGNANNGIATFTGDSTVTGNIGNTNAL 800

Qy        264 DAVNKA-------GWRMKTTTAN--------------------GQTGQADK-FETVTSGT 295
```

```
              ||         |  :|: | |                        || |:    | |  :
Db        801 ATVNVGAGLLRVQGGVVKSNTINLTDNASAVTFTNPVVVTGAIDNTGNANNGIVTFTGDS 860

Qy        296 KVTFASGN-GTTATVS--------------------KDDQGNITVKYDVNVGDALNVNQL 334
                ||   ||    ||:|                  |:   :|    | |  |::
Db        861 TVTGNIGNTNALATISVGAGKATLGGAIIKATTTKLTDNASAVTFTNPVVVTGAID---- 916

Qy        335 QNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEI--TRNGKNIDIATSM 392
              |:| | ::  |  :    ::||:  :      |||: ||  ::   : :  |||
Db        917 -NTG-NANNGIVTFTGDSTVTGNIGNTNAL--ATVNVGAGVTLQAGGSLDANNIDFGARS 972

Qy        393 TPQFSSVSLGAGADAP-----TLSVDDEGALNVGSK------------------------ 423
              | :|:    | |   |      ::  :   ||| :|
Db        973 TLEFNGPLDGGGNAIPYYFKGAIANGNNAILNVNTKLLTAYHLTIGTVAEINIGAGNLFA 1032

Qy        424 -DANK-PVRITN--------------------------------VAPGVKEGDV------ 443
               ||:   | | |                               |||| || |
Db       1033 IDASAGDVTILNAQDIHFRALDSALVLSNLTGVGVNNILLAADLVAPGVDEGTVVFDGGV 1092

Qy        444 ------TNVAQLKGVAQNLNNRIDNVNGNA--------RAGIAQAIATAGLVQAYLPGK- 488
                   :|||   | |:|    | :| ||            |   :   |:    :
Db       1093 NGLNIGSNVA---GAARN----IGDVGGNKFNTLLIYNAVTITDDVNLEGIQNVLINNNA 1145

Qy        489 ---SMMAIGGGT-YLGEAGYAI----GYSSISAGG-------NWIIKGTASGNSR 528
                | |   ||  : :| | |     |  :| ||          :|   :||| |
Db       1146 DFTSSTAFNAGTIQINDATYTIDANNGNLNIPAGNIKFAHADAQLILQNSSGNDR 1200


RESULT 7
OMPB_RICJA
ID   OMPB_RICJA     STANDARD;      PRT;  1656 AA.
AC   O06653;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Outer membrane protein B precursor (168 kDa surface-layer protein)
DE   (Surface protein antigen) (Cell surface antigen 5) (Sca5) (rOmpB)
DE   (rOmp B) [Contains: 120 kDa surface-exposed protein (Surface protein
DE   antigen) (120 kDa outer membrane protein ompB); 32 kDa beta peptide].
GN   OMPB.
OS   Rickettsia japonica.
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC   Rickettsiaceae; Rickettsieae; Rickettsia.
OX   NCBI_TaxID=35790;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=YH;
RA   Uchiyama T.;
RT   "Sequencing of the gene encoding the protein rOmp B of Rickettsia
RT   japonica.";
RL   Submitted (MAY-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: THE 120 KDA SURFACE-EXPOSED PROTEIN IS A MAJOR
CC       STRUCTURAL PROTEIN WHICH MAY PLAY A ROLE AS A RICKETTSIAL
CC       VIRULENCE FACTOR AND/OR IMMUNOGEN DURING INFECTION (BY
CC       SIMILARITY).
CC   -!- FUNCTION: THE 32 KDA BETA PEPTIDE MAY SERVE AS A MEMBRANE ANCHOR
CC       (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: CELL WALL. THIS BACTERIUM IS COVERED BY A S-
CC       LAYER WITH HEXAGONAL SYMMETRY.
CC   -!- SIMILARITY: BELONGS TO THE RICKETTSIAE OMPA/OMPB FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AB003681; BAA20138.1; -.
DR   InterPro; IPR006315; Autotransport.
DR   InterPro; IPR005546; Autotransporter.
DR   Pfam; PF03797; Autotransporter; 1.
```

```
DR   TIGRFAMs; TIGR01414; autotrans_barl; 2.
KW   Antigen; S-layer; Cell wall.
FT   CHAIN         1   1338       120 kDa SURFACE-EXPOSED PROTEIN.
FT   CHAIN      1339   1656       32 kDa BETA PEPTIDE.
FT   DOMAIN      528    533       POLY-GLY.
SQ   SEQUENCE   1656 AA;  168097 MW;  3132A69C9DD5999F CRC64;

  Query Match             6.8%;  Score 186;  DB 1;  Length 1656;
  Best Local Similarity   22.7%;  Pred. No. 0.071;
  Matches  136;  Conservative   79;  Mismatches  229;  Indels  154;  Gaps   33;

Qy         15 SVVGSIQASMEGSVELETISLSMTNDSKEFVDPYIVVTLKAGDNLKIKQNTNENTNASSF 74
              :| ||:  :::| :|| |:::              :   ||    :    |    |: |
Db        287 NVTGSLGGNLKGIIELNTVAINGQ-----------LIANAGPANAVIGTNNGAGRAAGF 334

Qy         75 TYSLKKDLTGLINVETEKLSFGANGKKVNIISDTKGLNFAKETAGTNGDTTV-HLNGI-- 131
               |:                |||   |      :   ::|  ||     |:  :
Db        335 VVSVD----------------NGKAATIDGQVYAKDMVIQSANANGQVNFRHIVDVGI 376

Qy        132 -GSTLTDMLLNTGATTNVTNDNVTDDEKKRA-ASVKDVLNAGWNIKG--VKPGTTASDNV 187
               |:|      :  | |  :|   ||     |  :| | :    |  |    || ||
Db        377 DGTTAFKTAASIVAITQNSNFGTTDFGNLAAQVTVPDTMTLTGNFTGDANNPGNTAG--- 433

Qy        188 DFVRTY---DTVEFLSADTKTTTVN----VESKDNG-------KKTEVKIGAKTSVIKEK 233
               | |:    |:   |||      |    :|:   |           |:::|   || |
Db        434 --VITFAANGTLASASADANVAVTNNITAIEASGVGVVQLSGTHTAELRLGNAGSVFKLA 491

Qy        234 DGKLVTGK--------GKGENGSSTDEGEGLVTAKEVIDAVNKAGWR--MKTTTANGQTG 283
              || :: ||        |    |: | :|   :|     |  |  |        | ||
Db        492 DGTVINGKVNQTVLVGGVLAAGAITLDGSATITG----DIGNGGGGAALQSITLAN---- 543

Qy        284 QADKFETVT---------SGTKVTFASGNGTTATVSKDDQGNITVKYDV-------NVGD 327
               |  :|:|         :|  : | :  ||     |   | || |  |:         | |
Db        544 --DATKTLTLGGANIISANGGTINFQANGGTIKLTS--TQNNIVVDCDLAIATDQTGVVD 599

Qy        328 ALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSK---GKMDETVN-----IN---AGNN 376
              | ::        |  :  ::|: | : : | :  :   |    |:|     ||    |||
Db        600 ASSLT-------NAQTLTISGTIGIIGANNTTLGQFNIGSSKTTLNGGNVAINELVIGNN 652

Qy        377 --IEITRNGKNIDIATSMTPQ----FSSV-----SLGAGADAPTLSVDDEGALNVGSKDA 425
               ::   |   |   |:   |    |: |     :| || :  : : :    :| ||| |
Db        653 GSVQFAHNTYLITRTTNAAGQGKIIFNPVVNNNTTLAAGTNLGS-AANPLAEINFGSKGA 711

Qy        426 NKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVNGNARAGIAQAIATAGLVQAYL 485
                  : ||  ||     ||:        |: : : |  |              :|   :
Db        712 RADT-VLNVGEGVNL-YATNITTTDA---NVGSFVFNAGGK-----------NIVSGTV 754

Qy        486 PGK-----SMMAIGGGT---YLGEAGYAIGYSSISAGGNWIIKGTASGNSRGHFGASA 535
               |:     : :|:  ||   :|| | :  | ::|:|     |    |||    | |||
Db        755 GGQQGNKFNTVALDNGTTVKFLGNATFN-GNTTIAANSTLQI----SGNYTADFIASA 807


RESULT 8
BIGA_SALTY
ID   BIGA_SALTY     STANDARD;      PRT;  1953 AA.
AC   P25927; P25928; Q9XCQ3;
DT   01-MAY-1992 (Rel. 22, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Putative surface-exposed virulence protein bigA precursor.
GN   BIGA OR STM3478.
OS   Salmonella typhimurium.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Salmonella.
OX   NCBI_TaxID=602;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 14028;
RA   Stojiljkovic I., Valentine P., Heffron F.;
RT   "Salmonella typhimurium rhs homolog.";
RL   Submitted (MAR-1999) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX   MEDLINE=21534948; PubMed=11677609;
RA   McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA   Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA   Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA   Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA   Waterston R., Wilson R.K.;
RT   "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT   LT2.";
RL   Nature 413:852-856(2001).
RN   [3]
RP   SEQUENCE OF 1-765 FROM N.A.
RC   STRAIN=LT2;
RX   MEDLINE=91100301; PubMed=1987123;
RA   Wu J.Y., Siegel L.M., Kredich N.M.;
RT   "High-level expression of Escherichia coli NADPH-sulfite reductase:
RT   requirement for a cloned cysG plasmid to overcome limiting siroheme
RT   cofactor.";
RL   J. Bacteriol. 173:325-333(1991).
CC   -!- CAUTION: Ref.3 sequence differs from that shown due to frameshifts
CC       in positions 414 and 732.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF133696; AAD39458.1; -.
DR   EMBL; AE008859; AAL22340.1; -.
DR   EMBL; M64606; AAA27042.1; ALT_FRAME.
DR   EMBL; M64606; AAA27043.1; ALT_FRAME.
DR   StyGene; SG10437; bigA.
KW   Virulence; Repeat; Signal; Complete proteome.
FT   SIGNAL        1     27       POTENTIAL.
FT   CHAIN        28   1953       PUTATIVE SURFACE-EXPOSED VIRULENCE
FT                                PROTEIN BIGA.
FT   DOMAIN      101    252       15 X 11 AA TANDEM REPEATS.
FT   REPEAT      101    103       1 (INCOMPLETE).
FT   REPEAT      104    113       2 (INCOMPLETE).
FT   REPEAT      114    122       3 (INCOMPLETE).
FT   REPEAT      123    133       4.
FT   REPEAT      134    144       5.
FT   REPEAT      145    155       6.
FT   REPEAT      156    166       7.
FT   REPEAT      167    177       8.
FT   REPEAT      178    188       9.
FT   REPEAT      189    199       10.
FT   REPEAT      200    210       11.
FT   REPEAT      211    221       12.
FT   REPEAT      222    232       13.
FT   REPEAT      233    243       14.
FT   REPEAT      244    252       15 (INCOMPLETE).
FT   CONFLICT    207    207       D -> DRGDDDVTPPDD (IN REF. 1).
FT   CONFLICT    514    514       A -> R (IN REF. 3).
FT   CONFLICT   1698   1698       D -> N (IN REF. 1).
FT   CONFLICT   1795   1798       QYLE -> ITLQ (IN REF. 1).
FT   CONFLICT   1836   1837       SA -> T (IN REF. 1).
SQ   SEQUENCE   1953 AA;  200150 MW;  611B3F1C954D91AE CRC64;

  Query Match             6.7%;  Score 184.5;  DB 1;  Length 1953;
  Best Local Similarity   24.0%;  Pred. No. 0.1;
  Matches  150;  Conservative   77;  Mismatches  201;  Indels  197;  Gaps   39;

Qy         48 YIVVTLKAGDNLKIKQNTNENTNASSFTYSLKKDLTGLI-----NVET----EKLSFGAN 98
              |:::   | ||  :    ::  |:        :| ||:|      | |      :: ||
Db        338 YLII---ADDNTTVISGDDQAHNSDRGMDISGQDRTGVIISGDRTVNTLTGDSSVTDGAT 394

Qy         99 GKKV------NIISD-------TKGL--------NFAKETAGTNGDT------------- 124
              |  :      | ||        |  |        |||  : | : | |
Db        395 GMVISGDGTTNTISGHSTVDNATGALISGNGTTTNFAGDIAVSGGGTAIIIDGDNATIKN 454
```

```
Qy        125 --TVHLNGIGSTLT------------DMLLNTGAT-TNVTNDNVTDDEKKRAASVKDVL 168
Db        455 TGTSDISGAGSTGTVIDGNNARVNNDGDMTITDGGTGGHITGDNVVID------------ 502

Qy        169 NAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDN------GKKTEVK 222
Db        503 NAG--------STTVS--------------GAD--ATALYIEG-DNALVINEGNQT-IS 535

Qy        223 IGAKTSVIKEKDGK-------LVTGKG------KGENGSSTDEGEGLVT--AKEVI--DA 265
Db        536 GGAVGTRIDGDDAHTTNTGDIAVDGAGSAAVIINGDNGSLTQAGDLLVTDGAMGIITYGT 595

Qy        266 VNKAGWRMKTTTANGQTGQADKFETVTSGTKVTF-------ASGNGTTATVSKDDQGNIT 318
Db        596 GNEA-----KNTGNATVRDADSVGFVVAGEKNTFKNKGDIDVSLNGTGALVS-GDMSQVT 649

Qy        319 VKYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKV-ISGNVSPS---------KGKMDET 368
Db        650 LDGDINV---VSVQDSEGVFSSATGVSVSGDSNAVDITGNVNISADYGQDDLAAGAPPLT 706

Qy        369 VNINAGNNIEITRNGK-NI---DIATSMTPQFSSVSLGAGADAPTLSVDDEGALNVGSKD 424
Db        707 GVVVGGNGNTVTLNGALNIDDNDLSATGGQYLDVVGLSVTGDDNDVEID--GGINI--TH 762

Qy        425 ANKPVRITNVAPGVKEGDVTNVAQLKG---VAQNLNNRIDN---VNGN---ARAGIAQAI 475
Db        763 SEDPLDGTS-------ADITGIS-VSGNSTVTLNGHSTIDTNTVVGGHVVLARVNNGGSL 814

Qy        476 -----ATAGLVQAYLP------GKSMMAIGGGTYLGEAG-------YAI----GYSSISA 513
Db        815 ILGDDSVVDVNVSYIPTGYYTYNALLMADGEGTSIENKGDITSHGVYSVIRADNGSEVSN 874

Qy        514 GGNWIIKGTASGNSRGHFGASASVG 538
Db        875 SGDILVYATSSNSSEDRAAITRASG 899
```

```
RESULT 9
SLAP_CAMFE
ID   SLAP_CAMFE     STANDARD;      PRT;   933 AA.
AC   P35827;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   01-OCT-1994 (Rel. 30, Last annotation update)
DE   S-layer protein (Surface array protein) (SAP).
GN   SAPA.
OS   Campylobacter fetus.
OC   Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
OC   Campylobacteraceae; Campylobacter.
OX   NCBI_TaxID=196;
RN   [1]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 1-20.
RC   STRAIN=84-32 / 23D;
RX   MEDLINE=90354448; PubMed=2387868;
RA   Blaser M.J., Gotschlich E.C.;
RT   "Surface array protein of Campylobacter fetus. Cloning and gene
RT   structure.";
RL   J. Biol. Chem. 265:14529-14535(1990).
RN   [2]
RP   ERRATUM.
RX   MEDLINE=91035477; PubMed=2229082;
RA   Blaser M.J., Gotschlich E.C.;
RL   J. Biol. Chem. 265:19372-19372(1990).
CC   -!- FUNCTION: THE S-LAYER IS A PARACRYSTALLINE MONO-LAYERED ASSEMBLY
CC       OF PROTEINS WHICH COAT THE SURFACE OF BACTERIA. THIS PROTEIN IS
CC       CRITICAL FOR VIRULENCE.
CC   -!- SUBCELLULAR LOCATION: CELL WALL. THIS BACTERIUM IS COVERED BY A
CC       S-LAYER WITH HEXAGONAL SYMMETRY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; J05577; AAA23032.1; -.
KW   Cell wall; S-layer.
SQ   SEQUENCE   933 AA;  96757 MW;  F88C729B4BA5B1E9 CRC64;

  Query Match             6.6%;  Score 180;  DB 1;  Length 933;
  Best Local Similarity   22.8%;  Pred. No. 0.071;
  Matches  133;  Conservative   69;  Mismatches  231;  Indels  150;  Gaps   23;

Qy          1 TDEDEEEELESVQRSVVGSIQASMEGSVELETISLS-MTNDSKEFVDPYIVVTLKAGDNL 59
Db        310 TDASKSVNVETT--GTITAFTAAGTGKVDVVAGKISALTADSRTSV-----------NL 355

Qy         60 KIKQNTNENTNASSFT-YSLK----KDLTGLINVETEKLSFGANGKKVNIISDTKGLNFA 114
Db        356 TATNDTITLTSANAATSVNLKQRQAKDAT-----ITSAMQQKYNNRRNRIATITSAT--A 408

Qy        115 KETAGTNGDTTVHLNGIGSTLTDMLLNTGA----------TTNVTNDNVTDDEK----- 158
Db        409 VENLTVKHATNVALNGGMDKLATVTLDNAALTAAIDIKSASTLNLINSSVNGPKHLYSSK 468

Qy        159 ------KRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVES 212
Db        469 RRYCKFKRAAA-----------KVKLNTTA----------------ATDQTVTLKANA 499

Qy        213 KDNGKKTEVKIGAKTSVIKEKDGKLVTGKGKGENGSSTDEGEGLVTAKEVIDAVNKAGWR 272
Db        500 TDNSLEFDSATAKTTSVTASGSGKTLVIKGA--------EVETLVN-------IDTTAFN 544

Qy        273 MKTTTANGQTGQADKF--ETVTSGTKVTFASGNGTTATVSK--DDQGNITVKYDVNVGDA 328
Db        545 ALQSVSFGKTGQGGKFSVKTGTGDDKIEFV---GTTLTEGSVIDAPGNDTIAMKSAALTS 601

Qy        329 LNVNQLQNSGWNLDSKAVAG---SSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKN 385
Db        602 ANFTMIKNIENVAISDAVATADLSSSAFKNSVIITTKEAADTTLTINKDQVI-------N 654

Qy        386 IDIATSMTPQFSSVSLGAGADAPTLSVDDEGALNVGSKDANKPVRITNVAPGVKEGDVTN 445
Db        655 FTAADAGSVKLITVKLN---DVTALMIVVKIVLDAAAKD-------TNIALGTAAADKAL 704

Qy        446 VAQLKGVAQNLNNRIDNVNGNARAGIAQAIATAGLVQAYLPGKSMMAIGG---------- 495
Db        705 V---------IDTGIETLNITSLVKATSPETTANTVNAKLTDVTSIIIDGMQITLGHAGT 755

Qy        496 -GTYLG----------EAGYAIGYSSISAGGNWIIKGTASGNS 527
Db        756 AGTDYSKVSMIDASALKAGLTFDASAITLGANATIKGGSGADS 798
```

```
RESULT 10
ALYS_ENTFA
ID   ALYS_ENTFA     STANDARD;      PRT;   737 AA.
AC   P37710;
DT   01-OCT-1994 (Rel. 30, Created)
DT   15-SEP-2003 (Rel. 42, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Autolysin precursor (EC 3.2.1.-) (Peptidoglycan hydrolase)
DE   (Beta-glycosidase).
GN   EF0799.
OS   Enterococcus faecalis (Streptococcus faecalis).
OC   Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.
OX   NCBI_TaxID=1351;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91358349; PubMed=1679432;
RA   Beliveau C., Potvin C., Trudel J., Asselin A., Bellemare G.;
RT   "Cloning, sequencing, and expression in Escherichia coli of a
RT   Streptococcus faecalis autolysin.";
```

```
RL   J. Bacteriol. 173:5619-5623(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=V583 / ATCC 700802;
RX   MEDLINE=22550857; PubMed=12663927;
RA   Paulsen I.T., Banerjei L., Myers G.S.A., Nelson K.E., Seshadri R.,
RA   Read T.D., Fouts D.E., Eisen J.A., Gill S.R., Heidelberg J.F.,
RA   Tettelin H., Dodson R.J., Umayam L., Brinkac L., Beanan M.,
RA   Daugherty S., DeBoy R.T., Durkin S., Kolonay J., Madupu R., Nelson W.,
RA   Vamathevan J., Tran B., Upton J., Hansen T., Shetty J., Khouri H.,
RA   Utterback T., Radune D., Ketchum K.A., Dougherty B.A., Fraser C.M.;
RT   "Role of mobile DNA in the evolution of vancomycin-resistant
RT   Enterococcus faecalis.";
RL   Science 299:2071-2074(2003).
CC   -!- FUNCTION: Hydrolyzes the cell wall of E.faecalis and
CC       M.lysodeikticus. May play an important role in cell wall growth
CC       and cell separation.
CC   -!- SUBCELLULAR LOCATION: Secreted (Probable).
CC   -!- DOMAIN: LYSM REPEATS ARE THOUGHT TO BE INVOLVED IN PEPTIDOGLYCAN
CC       BINDING.
CC   -!- SIMILARITY: BELONGS TO FAMILY 73 OF GLYCOSYL HYDROLASES.
CC   -!- SIMILARITY: Contains 6 LysM repeats.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M58002; AAA67325.1; -.
DR   EMBL; AE016949; AAO80613.1; -.
DR   PIR; A38109; A38109.
DR   TIGR; EF0799; -.
DR   InterPro; IPR002901; Amidase_4.
DR   InterPro; IPR002482; LysM.
DR   Pfam; PF01832; Amidase_4; 1.
DR   Pfam; PF01476; LysM; 5.
DR   SMART; SM00257; LysM; 6.
DR   SMART; SM00047; LYZ2; 1.
KW   Hydrolase; Glycosidase; Bacteriolytic enzyme; Cell wall;
KW   Cell division; Septation; Repeat; Signal.
FT   SIGNAL        1     53       POTENTIAL.
FT   CHAIN        54    737       AUTOLYSIN.
FT   REPEAT      363    405       LYSM 1.
FT   REPEAT      431    473       LYSM 2.
FT   REPEAT      499    541       LYSM 3.
FT   REPEAT      567    609       LYSM 4.
FT   REPEAT      633    675       LYSM 5.
FT   REPEAT      695    737       LYSM 6.
FT   CONFLICT     85     85       T -> I (IN REF. 1).
FT   CONFLICT    118    118       A -> V (IN REF. 1).
FT   CONFLICT    143    143       A -> T (IN REF. 1).
FT   CONFLICT    417    417       S -> N (IN REF. 1).
FT   CONFLICT    449    449       S -> T (IN REF. 1).
FT   CONFLICT    476    476       A -> T (IN REF. 1).
FT   CONFLICT    484    484       N -> S (IN REF. 1).
FT   CONFLICT    567    632       MISSING (IN REF. 1).
SQ   SEQUENCE   737 AA;  77025 MW;  ABB16BD506AC7507 CRC64;

  Query Match             6.6%;  Score 179.5;  DB 1;  Length 737;
  Best Local Similarity   23.0%;  Pred. No. 0.057;
  Matches  111;  Conservative   55;  Mismatches  208;  Indels  109;  Gaps   21;

Qy         94 SFGANGKKVNIISDTKGLNFA----KETAGTNGDTTVHLNGIGSTLTDMLLNTGATTNVT 149
              ||  |   :   |   |: :     |    :  | |  |   |   ||   |       :|
Db        268 SFQDNAHVLKTTSFQAGVYYYAGAWKSNTSSYRDATAWLT--GRYATDPSYNAKLNNVIT 325

Qy        150 NDNVTDDEKKRAASVKDVLNAGWNIKG--VKPGTTASDNVDFVRTY------DTVEFLSA 201
              |:|           |  ::| |  |  | |||  |:|     ||      ||:  ::|
Db        326 AYNLTQ---------YDTPSSGGNTGGGTVNPGTGGSNNQSGTNTYYTVKSGDTLNKIAA 376

Qy        202 DTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKGKGENGSSTDE-GEGLVTAK 260
                  :  |: | :      : :| |  | |   |   || | |  ||: :: |
Db        377 QYGVSVANLRSWNGISGDLIFVGQKLIVKKGASGN--TG-GSGSGGSNNNQSGTNTYYTV 433

Qy        261 EVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTKVTF---ASGN--GTTATVSKDDQG 315
              :  | :||   :   : || ::      : :  | |:     ||||  |:    | ::|
Db        434 KSGDTLNKIAAQYGVSVANLRSWNGISGDLIFVGQKLIVKKGASGNTGGSNNGGSNNNQS 493

Qy        316 NITVKYDVNVGDALN-------VNQLQNSGWNLDSKAVAGSSGKVI-----SGNVSPSKG 363
                  | :  || ||       |:          ||  |  :  :  |:|     |||   |
Db        494 GTNTYYTIKSGDTLNKIAAQYGVSVANLRSWNGISGDLIFAGQKIIVKKGTSGNTGGSS- 552

Qy        364 KMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSV--DDEGAL--- 418
                    | |:|    ::| |           : :|   |     |    :  |:|
Db        553 --------NGGSN--NNQSGTNTYYTIKSGDTLNKISAQFGVSVANLQAWNNISGSLIFA 602

Qy        419 ----------NVGSKDANKPVRITNVAPG------VKEGDVTNVAQLKGVAQNLNNRIDN 462
                       | || : |||   ||   |      :| ||  |
Db        603 GQKIIVKKGANSGSTNTNKP---TNNGGGATTSYTIKSGDTLN----------------K 643

Qy        463 VNGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYLGEAGYAIGYSSISAGGNWIIKGT 522
              ::      :|   :  |:       |  :   | | : : |        ||||    :
Db        644 ISAQFGVSVANLRSWNGI-------KGDLIFAGQTIIVKKG-------ASAGGNASSTNS 689

Qy        523 ASG 525
              |||
Db        690 ASG 692


RESULT 11
HLYA_SERMA
ID   HLYA_SERMA     STANDARD;      PRT;  1608 AA.
AC   P15320;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   01-NOV-1990 (Rel. 16, Last annotation update)
DE   Hemolysin precursor.
GN   SHLA.
OS   Serratia marcescens.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Serratia.
OX   NCBI_TaxID=615;
RN   [1]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 31-40.
RC   STRAIN=SN8;
RX   MEDLINE=88257037; PubMed=3290200;
RA   Poole K., Schiebel E., Braun V.;
RT   "Molecular characterization of the hemolysin determinant of Serratia
RT   marcescens.";
RL   J. Bacteriol. 170:3177-3188(1988).
CC   -!- FUNCTION: BACTERIAL HEMOLYSINS ARE EXOTOXINS THAT ATTACK BLOOD
CC       CELL MEMBRANES AND CAUSE CELL RUPTURE BY MECHANISMS NOT CLEARLY
CC       DEFINED.
CC   -!- FUNCTION: CELL-BOUND HEMOLYSIN, WHICH RELEASES HEME-IRON FROM
CC       ERYTHROCYTES BY INTERACTION WITH THE ERYTHROCYTE MEMBRANE. SHLA
CC       REQUIRES SHLB FUNCTION.
CC   -!- SUBCELLULAR LOCATION: Outer membrane.
CC   -!- SIMILARITY: TO P.MIRABILIS HEMOLYSIN (HPMA).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M22618; AAA50323.1; -.
DR   PIR; A28182; A28182.
KW   Hemolysis; Toxin; Outer membrane; Signal.
FT   SIGNAL        1     30
FT   CHAIN        31   1608       HEMOLYSIN.
SQ   SEQUENCE   1608 AA;  165078 MW;  D669B476FE7DAD51 CRC64;
```

```
  Query Match             6.5%;  Score 179;  DB 1;  Length 1608;
  Best Local Similarity   22.1%;  Pred. No. 0.15;
  Matches  152;  Conservative   81;  Mismatches  232;  Indels  224;  Gaps   34;

Qy          2 DEDEEEELESVQRSVVGSIQAS-MEGSVELETISLSMTNDSKEFVDPYIVVTLKAGDNL- 59
              ::|:  |    ||:   |::      | | |::  |       |       :||||| |:
Db        449 EKDKSSE-RGYQRNHTSSLRTGRWSNSDESESLKASELRSEGE-------LTLKAGRNVS 500

Qy         60 ------------------KIKQNTNENTNASSFTYSLKKDLTGLINVETEKLSFGANG-- 99
                              :|:    :  || :             |  :| |:|  |
Db        501 TQGAKVHAQRDLTIDADNQIQVGVQKTANAKA-------------VRDDKTSWGGIGGG 546

Qy        100 -KKVNIISDTKGLNFAKETAGTNGDTTVHLNG-IGSTLTDMLL--NTGATTNVTNDNVTD 155
               |  |  |: : :: | |   |:|  |: |||  | |:|        |     |:  :
Db        547 DNKNN--SNRREISHASEL--TSGG-TLRLNGQQGVTITGSKARGQKGGEVTATHGGLRI 601

Qy        156 DEKKRAASVKDVLNA----GWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVE 211
               |     ::  | ::|     :||      |::|   |       |   | :||  | |: :
Db        602 D--NALSTTVDKIDART-GTAFNI------TSSSHKADNSYQSSTASELKSDTNLTLVSHK 653

Qy        212 SKD-NGKKT----EVKIGAKTSVIKEKDGKLVTGKGKGENGSSTDEGEGLVT----AKEV 262
               |  | :     |: : :||  |  |           |   : || :  :|    |||
Db        654 DADVIGSQVASGGGELSVESKTGNINVK---------AAERQQNIDEQKTALTVNGYAKEA 704

Qy        263 IDAVNKAGWRM-------KTT----------------------TANGQTGQADKFETVTS 293
               |   :|| |:       |||                      | :|    |||  :   |
Db        705 GDKQYRAGLRIEHTRDSEKTTRTENSASSLSGGSVKLKAEKDVTFSGSKLVADKGDASVS 764

Qy        294 GTKVTFASGNGTTA-------------------------------------------TVSK 311
               | ||:| : :  ||                                          |
Db        765 GNKVSFLAADDKTASNTEQTKIGGGFYYTGGIDKLGSGVEAGYENNKTQAQSSKAITSGS 824

Qy        312 DDQGNITVKYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNI 371
               | :||:|    :|  | |     |:|      :  ||      :   | :  | |  |||
Db        825 DVKGNLT----INARDKLTQQGAQHSVGGAYQENAAGVDHLAAADTASTTTKTDVGVNI 880

Qy        372 NAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEG-------ALNVGSKD 424
                        | |:| :    |   :|   |  || |  ::| |         |::|::
Db        881 -----------GANVDYSAVTRPVERAVGKAAKLDA-TGVINDIGGIGAPNVGLDIGAQG 928

Qy        425 ANKPVRITN---VAPGVKEGDV------------TNVAQLKGVAQNLN----------NR 459
               :  | ::   |   |: | :            |     || | ||           ||
Db        929 GSSEKRSSSSQAVVSSVQAGSIDINAKGEVRDQGTQYQASKG-AVNLTADSHRSEAAANR 987

Qy        460 IDNVNGNARAGIAQAIATAGLVQAYLPGKSMMAIG----GGTYL--GEAGYAIGYSSISA 513
               |  : : |        :|| |: |    | : :     |||      |  |:  |  :|
Db        988 QDEQSRDTR-------GSAG-VRVYTTTGSDLTVDAKGEGGTQRSNSSASQAVTGSIDAA 1039

Qy        514 GG-------NWIIKGTASGNSRGHFGASA 535
                |       : | :|||    ||    :|
Db       1040 NGINVNVKKDAIYQGTALNGGRGKTAVNA 1068
```

```
RESULT 12
OMPB_RICTY
ID   OMPB_RICTY     STANDARD;      PRT;  1645 AA.
AC   P96989;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Outer membrane protein B precursor (168 kDa surface-layer protein)
DE   (Surface protein antigen) (Cell surface antigen 5) (Sca5) (rOmpB)
DE   (rOmp B) [Contains: 120 kDa surface-exposed protein (Surface protein
DE   antigen) (120 kDa outer membrane protein ompB); 32 kDa beta peptide].
GN   OMPB OR SLP.
OS   Rickettsia typhi.
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC   Rickettsiaceae; Rickettsieae; Rickettsia.
OX   NCBI_TaxID=785;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Wilmington;
RX   MEDLINE=94040787; PubMed=8224886;
RA   Hahn M.-J., Kim K.-K., Kim I., Chang W.-H.;
RT   "Cloning and sequence analysis of the gene encoding the crystalline
RT   surface layer protein of Rickettsia typhi.";
RL   Gene 133:129-133(1993).
RN   [2]
RP   PARTIAL SEQUENCE.
RC   STRAIN=Wilmington;
RX   MEDLINE=92114896; PubMed=1370573;
RA   Ching W.M., Carl M., Dasch G.A.;
RT   "Mapping of monoclonal antibody binding sites on CNBr fragments of
RT   the S-layer protein antigens of Rickettsia typhi and Rickettsia
RT   prowazekii.";
RL   Mol. Immunol. 29:95-105(1992).
RN   [3]
RP   IDENTIFICATION OF CLEAVAGE SITE.
RX   MEDLINE=92104668; PubMed=1729180;
RA   Hackstadt T., Messer R., Cieplak W., Peacock M.G.;
RT   "Evidence for proteolytic cleavage of the 120-kilodalton outer
RT   membrane protein of rickettsiae: identification of an avirulent
RT   mutant deficient in processing.";
RL   Infect. Immun. 60:159-165(1992).
CC   -!- FUNCTION: THE 120 kDa SURFACE-EXPOSED PROTEIN IS A MAJOR
CC       STRUCTURAL PROTEIN WHICH MAY PLAY A ROLE AS A RICKETTSIAL
CC       VIRULENCE FACTOR AND/OR IMMUNOGEN DURING INFECTION.
CC   -!- FUNCTION: THE 32 kDa BETA PEPTIDE MAY SERVE AS A MEMBRANE ANCHOR.
CC   -!- SUBCELLULAR LOCATION: CELL WALL. THIS BACTERIUM IS COVERED BY A
CC       S-LAYER WITH HEXAGONAL SYMMETRY.
CC   -!- SIMILARITY: BELONGS TO THE RICKETTSIAE OMPA/OMPB FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; L04661; AAB48987.1; -.
DR   PIR; JN0896; JN0896.
DR   InterPro; IPR006315; Autotransport.
DR   InterPro; IPR005546; Autotransporter.
DR   Pfam; PF03797; Autotransporter; 1.
DR   TIGRFAMs; TIGR01414; autotrans_barl; 1.
KW   Antigen; S-layer; Transmembrane; Cell wall.
FT   CHAIN         1   1353       120 kDa SURFACE-EXPOSED PROTEIN.
FT   CHAIN      1354   1645       32 kDa BETA PEPTIDE.
FT   TRANSMEM   1415   1429       MEMBRANE ANCHOR (POTENTIAL).
FT   CONFLICT    657    657       H -> N (IN REF. 2).
FT   CONFLICT    842    842       V -> I (IN REF. 2).
FT   CONFLICT   1071   1071       G -> A (IN REF. 2).
FT   CONFLICT   1306   1306       G -> S (IN REF. 2).
SQ   SEQUENCE   1645 AA;  169698 MW;  0CB5641C7EB185EE CRC64;
```

```
  Query Match             6.5%;  Score 179;  DB 1;  Length 1645;
  Best Local Similarity   22.3%;  Pred. No. 0.16;
  Matches  146;  Conservative   67;  Mismatches  233;  Indels  208;  Gaps   33;

Qy         15 SVVGSIQASMEGSVELETISLSMTNDSKEFVDPYIVVTLKAGDNLKIKQNTNENTNASSF 74
              | |||| :   ||    :  :: | : :          |  ||| ||    |   | ::
Db        170 SGVGSIDFTAAPSV----LEFNLINPTTQ------EAPLTLGDNAKIVNGANGILNITN- 218

Qy         75 TYSLKKDLTGLINVETEKLSFGANGKKVNIISDTKGLNFAKETAGTN------GDTTVHL 128
                        | : |  :  :    | |   | | :|| |       |      |  |::
Db        219 ---------GFVKVSDKTFA----GIKTINIGDNQGLMFNTTPDAANALNLQGGGNTINF 265

Qy        129 NGIGSTLTDMLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGV-KPGTTA---- 183
              ||      |  |:       | |  |||              : | |:||| :  |||
Db        266 NGRDG--TGKLVLVSKNGNATEFNVTG-------------SLGGNLKGVIEFDTTAAAGK 310

Qy        184 -------------SDN-----VDFVRTYDTVEFLSADTKTTTVNVESKD--------NGK 217
                          :||       |: : |     : :  | :  | :||         |:
Db        311 LIANGGAANAVIGTDNGAGRAAGFIVSVD-----NGNAATISGQVYAKDIVIQSANAGGQ 365
```

```
Qy        218 KT-----EVKIGAKTSVIKEKDGKLVTGKGKGENGS--STDEG----------EGLVTAK 260
Db        366 VTFEHLVDVGLGGKTN-FKTADSKVII----TENASFGSTDFGNLAVQIVVPNNKILTGN 420

Qy        261 EVIDAVNK---AGWRMKTTTANGQ--TGQADKFETVTS-------GTKVTFASG------ 302
Db        421 FIGDAKNNGNTAG--VITFNANGTLVSGNTDPNIVVTNIKAIEVEGAGIVQLSGIHGAEL 478

Qy        303 -NGTTATVSKDDQGNITVKYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPS 361
Db        479 RLGNAGSIFKLADGTV-INGPVNQNPLVNNNALAAGSIQLDGSAII--TGDI--GNGAVN 533

Qy        362 KGKMDET----------------VNINAGNNIEITRNGKNI--------------DIAT 390
Db        534 AALQDITLANDASKILTLSGANIIGANAGGAIHFQANGGTIQLTSTQNNILVDFDLDVTT 593

Qy        391 SMTPQFSSVSL-------------GAGADAPTLSVDDEGALNVGSK------------- 423
Db        594 DQTGVVDASSLTNNQTLTINGSIGTIGANTKTL-----GRFNVGSSKTILNAGDVAINEL 648

Qy        424 --DANKPVRIT-NVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVNGNARAGIAQAIATAGL 480
Db        649 VMENDGSVHLTHNTYLITKTINAANQGKIIVAADPINTDTALADGTNLGSAESPLSNIHF 708

Qy        481 VQAYLPGKSMMAIGGGTYLGEAGYA---------IGYSSISAGGNWIIKGTASG 525
Db        709 ATKAANGDSILHIGKGVNL----YANNITTTDANVGSLHFRSGGTSIVSGTVGG 758
```

RESULT 13
AG43_ECOLI
ID AG43_ECOLI STANDARD; PRT; 1039 AA.
AC P39180; P75614; P76360; P97241; Q46771;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Antigen 43 precursor (AG43) (Fluffing protein).
GN FLU OR B2000.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=97251358; PubMed=9097040;
RA Itoh T., Aiba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K.,
RA Kasai H., Kimura S., Kitakawa M., Kitagawa M., Makino K., Miki T.,
RA Mizobuchi K., Mori H., Mori T., Motomura K., Nakade S., Nakamura Y.,
RA Nashimoto H., Nishio Y., Oshima T., Saito N., Sampei G., Seki Y.,
RA Sivasundaram S., Tagami H., Takeda J., Takemoto K., Wada C.,
RA Yamamoto Y., Horiuchi T.;
RT "A 460-kb DNA sequence of the Escherichia coli K-12 genome
RT corresponding to the 40.1-50.0 min region on the linkage map.";
RL DNA Res. 3:379-392(1996).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=ML 308-225;
RA Henderson I.R., Owen P.;
RL Submitted (JAN-1997) to the EMBL/GenBank/DDBJ databases.
RN [4]
RP PRELIMINARY SEQUENCE OF 53-78.
RC STRAIN=ML 308-225;
RX MEDLINE=89291704; PubMed=2661530;
RA Caffrey P., Owen P.;
RT "Purification and N-terminal sequence of the alpha subunit of antigen
RT 43, a unique protein complex associated with the outer membrane of
RT Escherichia coli.";
RL J. Bacteriol. 171:3634-3640(1989).
RN [5]
RP SEQUENCE OF 53-63.
RC STRAIN=K12 / EMG2;
RX MEDLINE=97443975; PubMed=9298646;
RA Link A.J., Robison K., Church G.M.;
RT "Comparing the predicted and observed properties of proteins encoded
RT in the genome of Escherichia coli K-12.";
RL Electrophoresis 18:1259-1313(1997).
RN [6]
RP GENE NAME.
RX MEDLINE=97257509; PubMed=9103983;
RA Henderson I.R., Meehan M., Owen P.;
RT "Antigen 43, a phase-variable bipartite outer membrane protein,
RT determines colony morphology and autoaggregation in Escherichia coli
RT K-12.";
RL FEMS Microbiol. Lett. 149:115-120(1997).
CC -!- FUNCTION: CONTROLS COLONY FORM VARIATION AND AUTOAGGREGATION. MAY
CC FUNCTION AS AN ADHESIN.
CC -!- SUBUNIT: HETERODIMER OF THE PERIPHERAL MEMBRANE PROTEIN (ALPHA
CC CHAIN) ANCHORED TO THE INTEGRAL OUTER MEMBRANE PROTEIN (BETA
CC CHAIN).
CC -!- SUBCELLULAR LOCATION: Outer membrane-associated.
CC -!- SIMILARITY: TO ADHESIN AIDI-I AND TO BORDETELLA PERTACTIN.
CC --

CC --
DR EMBL; AE000291; AAC75061.1; ALT_INIT.
DR EMBL; D90838; BAA15825.1; ALT_INIT.
DR EMBL; D90839; BAA15832.1; ALT_INIT.
DR EMBL; U24429; AAB47869.1; -.
DR HSSP; P07505; 1SRD.
DR EcoGene; EG12686; flu.
DR InterPro; IPR006315; Autotransport.
DR InterPro; IPR005546; Autotransporter.
DR InterPro; IPR004899; Pertactin.
DR Pfam; PF03797; Autotransporter; 1.
DR Pfam; PF03212; Pertactin; 1.
DR TIGRFAMs; TIGR01414; autotrans_barl; 1.
KW Outer membrane; Signal; Complete proteome.

FT	SIGNAL	1	52	
FT	CHAIN	53	551	ANTIGEN 43 ALPHA CHAIN.
FT	CHAIN	552	1039	ANTIGEN 43 BETA CHAIN.
FT	VARIANT	2	2	K -> N (IN STRAIN ML 308-225).
FT	VARIANT	41	42	SL -> FF (IN STRAIN ML 308-225).
FT	VARIANT	46	46	T -> K (IN STRAIN ML 308-225).
FT	VARIANT	157	157	W -> L (IN STRAIN ML 308-225).
FT	VARIANT	188	188	V -> F (IN STRAIN ML 308-225).
FT	VARIANT	303	305	ATN -> STI (IN STRAIN ML 308-225).
FT	VARIANT	320	320	A -> T (IN STRAIN ML 308-225).
FT	VARIANT	372	372	N -> Q (IN STRAIN ML 308-225).
FT	VARIANT	493	493	E -> V (IN STRAIN ML 308-225).
FT	VARIANT	497	497	S -> N (IN STRAIN ML 308-225).
FT	VARIANT	585	585	H -> Y (IN STRAIN ML 308-225).
FT	VARIANT	709	709	E -> K (IN STRAIN ML 308-225).
FT	VARIANT	721	721	M -> T (IN STRAIN ML 308-225).
FT	VARIANT	751	753	GHL -> SHF (IN STRAIN ML 308-225).
FT	VARIANT	803	803	S -> P (IN STRAIN ML 308-225).
FT	VARIANT	815	815	A -> V (IN STRAIN ML 308-225).
FT	VARIANT	824	824	C -> S (IN STRAIN ML 308-225).
FT	VARIANT	829	835	LNLVHTS -> MNLIYNA (IN STRAIN ML 308-

```
FT                              225).
FT   VARIANT     845    847       QGT -> LGA (IN STRAIN ML 308-225).
FT   VARIANT     855    855       S -> T (IN STRAIN ML 308-225).
FT   VARIANT     888    888       Q -> L (IN STRAIN ML 308-225).
FT   VARIANT    1025   1025       S -> I (IN STRAIN ML 308-225).
FT   CONFLICT     61     63       ETV -> TTT (IN REF. 5).
SQ   SEQUENCE   1039 AA;  106841 MW;  5170D647C8DEEBE0 CRC64;

  Query Match             6.5%;  Score 177;  DB 1;  Length 1039;
  Best Local Similarity   21.1%;  Pred. No. 0.11;
  Matches  120;  Conservative   74;  Mismatches  221;  Indels  153;  Gaps   27;

Qy         52 TLKAGDNLKIKQNTNENTNASSFTYSLKKDLTGLINVETEKLSFGANGKKVNIISDTKGL 111
              ||   ||  :   ||  | ::   |    :     |   : :  |    | : |   ||
Db         67 TLANHDNQIVFGTTNGMTISTGLEYGPDNE----ANTGGQWVQDGGTANKTTVTSG--GL 120

Qy        112 NFAKETAGTNGDTTVHLNGIGSTLTDMLLNTG------------ATTNVTNDNVTDDEK 158
                    |:  || :   | | :|    :||              ||  | ||
Db        121 Q-RVNPGGSVSDTVISAGG-GQSLQGRAVNTTLNGGEQWMHEGAIATGTVINDK------ 172

Qy        159 KRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRT---------YDTVEFLSADTKTTTV 208
                       || :  ||||| |:| |  | |           || :|:  |   ||:
Db        173 -----------GWQV--VKPGTVATDTV--VNTGAEGGPDAENGDTGQFVRGDAVRTTI 216

Qy        209 N------VESKDNGKKTEVKIGAKTSV-----------------------IKEKDGKL 237
              |      | ::    | |  |   :|                       :   ||
Db        217 NKNGRQIVRAEGTANTTVVYAGGDQTVHGHALDTTLNGGYQYVHNGGTASDTVVNSDGWQ 276

Qy        238 VTGKGKGENGSSTDEGEGLVTAKEVIDAVN---KAGWRMKTTTANGQT------------ 282
              :   | |  |::|   :| :       | |   | |  : |:||   |
Db        277 IVKNG-GVAGNTTVNQKGRLQVDAGGTATNVTLKQGGALVTSTAATVTGINRLGAFSVVE 335

Qy        283 GQADKFETVTSGTKVTFASGNGTTATVSKDDQGNITVKYDVNVGDALNVNQLQNSGWNLD 342
              |:||    : :| ::   :|:  | |   || | : |:   | | |  |: : | |
Db        336 GKADNV-VLENGGRLDVLTGHTATNT-RVDDGGTLDVR---NGGTATTVS-MGNGG---- 385

Qy        343 SKAVAGSSGKVISGNVSPSK-----GKMDETVNINAGNNIEITRNGKNIDIATSMTPQFS 397
                 ||  :|| |  |     |   : : :  |::  :       | ||  |
Db        386 --VLLADSGAAVSGTRSDGKAFSIGGGQADALMLEKGSSFTLNAG----DTATDTTVNGG 439

Qy        398 SVSLGAGADAPTLSVDDEGALNVGSKDANKPVRITNVAPGVKEGDVTNVAQLKGVAQNLN 457
              :   |  | | ::::   | :  |  |      |    ::|||    | |:| :   |
Db        440 LFTARGGTLAGTTTLNNGAILTLSGKTVN------NDTLTIREGD----ALLQGGSLTGN 489

Qy        458 NRIDNVNGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGT----YLGEAGYAIGYS---- 509
              ::  :|:   ::    |  |      :  : :  |     : | |:  :
Db        490 GSVEK-SGSGTLTVSNTTLTQKAVNL---NEGTLTLNDSTVTTDVIAQRGTALKLTGSTV 545

Qy        510 ----------SISAGGNWIIKGTASGNS 527
                       ::::|  | |   |:  |
Db        546 LNGAIDPTNVTLASGATWNIPDNATVQS 573


RESULT 14
YDEK_ECOLI
ID   YDEK_ECOLI     STANDARD;      PRT;  1325 AA.
AC   P32051; P76140; P77168;
DT   01-OCT-1993 (Rel. 27, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Hypothetical lipoprotein ydeK precursor (ORFT).
GN   YDEK OR ORFT OR B1510.
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12 / MG1655;
RX   MEDLINE=97426617; PubMed=9278503;
RA   Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA   Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA   Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA   Mau B., Shao Y.;
RT   "The complete genome sequence of Escherichia coli K-12.";
RL   Science 277:1453-1474(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12;
RX   MEDLINE=97251357; PubMed=9097039;
RA   Aiba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K., Itoh T.,
RA   Kasai H., Kashimoto K., Kimura S., Kitakawa M., Kitagawa M.,
RA   Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,
RA   Nakade S., Nakamura Y., Nashimoto H., Nishio Y., Oshima T., Saito N.,
RA   Sampei G., Seki Y., Sivasundaram S., Tagami H., Takeda J.,
RA   Takemoto K., Takeuchi Y., Wada C., Yamamoto Y., Horiuchi T.;
RT   "A 570-kb DNA sequence of the Escherichia coli K-12 genome
RT   corresponding to the 28.0-40.1 min region on the linkage map.";
RL   DNA Res. 3:363-377(1996).
RN   [3]
RP   SEQUENCE OF 595-1325 FROM N.A.
RX   MEDLINE=94100243; PubMed=8274505;
RA   Cartwright P.J., Timms M.W., Lithgow T., Hoej P.B., Hoogenraad N.J.;
RT   "An Escherichia coli gene showing a potential ancestral relationship
RT   to the genes for the mitochondrial import site proteins ISP42 and
RT   MOM38.";
RL   Biochim. Biophys. Acta 1153:345-347(1993).
CC   -!- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor
CC       (Potential).
CC   -!- SIMILARITY: TO E.COLI YFAL.
CC   -!- SIMILARITY: SOME, TO FUNGAL MITOCHONDRIAL IMPORT SITE PROTEINS
CC       ISP42 AND MOM38.
CC   -!- CAUTION: REF.3 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO A
CC       FRAMESHIFT IN POSITION 653.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; AE000248; AAC74583.1; -.
DR   EMBL; D90793; BAA15190.1; ALT_INIT.
DR   EMBL; D90794; BAA15197.1; ALT_INIT.
DR   EMBL; X73295; CAA51730.1; ALT_FRAME.
DR   PIR; A64905; A64905.
DR   EcoGene; EG11780; ydeK.
DR   PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
KW   Hypothetical protein; Membrane; Lipoprotein; Signal;
KW   Complete proteome.
FT   SIGNAL        1     18       POTENTIAL.
FT   CHAIN        19   1325       HYPOTHETICAL LIPOPROTEIN YDEK.
FT   LIPID        19     19       N-ACYL DIGLYCERIDE (POTENTIAL).
FT   CONFLICT    884    884       N -> K (IN REF. 3).
FT   CONFLICT   1317   1317       M -> S (IN REF. 3).
SQ   SEQUENCE   1325 AA;  136514 MW;  26A3A066FA19AD7D CRC64;

  Query Match             6.4%;  Score 175.5;  DB 1;  Length 1325;
  Best Local Similarity   23.4%;  Pred. No. 0.18;
  Matches  135;  Conservative   69;  Mismatches  237;  Indels  135;  Gaps   28;

Qy         16 VVGSIQASMEGSVELETISLSMTN-DSKEFVDPYI-------VVTLKAGDNLKIKQNTNE 67
              |: ::  |  |:  ||  :  :|| |:    | |:       |: : || |  :
Db         51 VLLALSGSASGA-SLEVDNDQITNIDTDVAYDAYLVGWYGTGVLNILAGGNASL------ 103

Qy         68 NTNASSFTYSLKKDLTGLINV--ETEKL-SFGANGKKVNI-ISDTKGLN----------F 113
                 |  ::      :|  | :||   | :|   | | : :|:  | |  ||          :
Db        104 -TTITTSVIGANEDSEGTVNVLGGTWRLYDSGNNARPLNVGQSGTGTLNIKQKGHVDGGY 162

Qy        114 AKETAGTNGDTTVHLNGIGSTLTDMLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWN 173
                 :  : | |  ||:: |  | ||  |   | :    : |:||
Db        163 LRLGSSTGGVGTVNVEGEDSVLTTELFEIG-SYGTGSLNITD----------------- 203

Qy        174 IKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEK 233
```

```
Db        204 -KG---------------YVTSSIVA------ILGYQAGSNGQVVVEKGG--EWLIKNN 238

Qy        234 DGKLVTGKGKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTS 293
Db        239 DSSIEFQIGNQGTGEATIREGGLVTAENTIIGGNATG----IGTLNVQ----DQDSVITV 290

Qy        294 GTKVTFASGNGTTATVSKDDQGNITVK-YD-VNVGDALN--VNQLQNSGWNL-------- 341
Db        291 RRLYNGYFGNG---TVNISNNGLINNKEYSLVGVQDGSHGVVNVTDKGHWNFLGTGEAFR 347

Qy        342 --------DSKAVAGSSGKVISGNVSPSKGKMDETVNINAGN------NIEITRNGKNID 387
Db        348 YIYIGDAGDGELNVSSEGKVDSGIITAG---MKET---GTGNITVKDKNSVITNLGTNLG 401

Qy        388 IATSMTPQFSSVSLGAGADAPTLSVDDEGALNV-----GSKDANKPVRITNVAPGVKEGD 442
Db        402 YDGHGEMNISNQGLVVSNGGSSLGYGETGVGNVSITTGGMWEVNKNVYTTIGVAGVGNLN 461

Qy        443 VTNVAQLKGVAQNLNNRIDNVNGNARAGIAQAIATAGLVQAYLP--GKSMMAIGGGTYLG 500
Db        462 ISDGG--KFVSQNITFLGDKASGIGTLNLMDATSSFDTVGINVGNFGSGIVNVSNGATLN 519

Qy        501 EAGYAI--------GYSSISAGGNWIIKGTASGNSR 528
Db        520 STGYGFIGGNASGKGIVNISTDSLWNLK-TSSTNAQ 554
```

```
RESULT 15
ICEN_XANCT
ID   ICEN_XANCT     STANDARD;      PRT;  1567 AA.
AC   P18127;
DT   01-NOV-1990 (Rel. 16, Created)
DT   01-NOV-1990 (Rel. 16, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Ice nucleation protein.
GN   INAX.
OS   Xanthomonas campestris (pv. translucens).
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC   Xanthomonadaceae; Xanthomonas.
OX   NCBI_TaxID=343;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=X56S;
RX   MEDLINE=91080859; PubMed=2259339;
RA   Zhao J., Orser C.S.;
RT   "Conserved repetition in the ice nucleation gene inaX from
RT   Xanthomonas campestris pv. translucens.";
RL   Mol. Gen. Genet. 223:163-166(1990).
CC   -!- FUNCTION: ICE NUCLEATION PROTEINS ENABLE BACTERIA TO NUCLEATE
CC       CRYSTALLIZATION IN SUPERCOOLED WATER.
CC   -!- SUBCELLULAR LOCATION: Outer membrane (By similarity).
CC   -!- DOMAIN: CONTAINS 153 IMPERFECT REPEATS OF THE CONSENSUS
CC       OCTAPEPTIDE A-G-Y-G-S-T-L-T; FURTHER ON A 16-RESIDUE AND A
CC       REGIONAL 48-RESIDUE PERIODICITY IS SUPERIMPOSED.
CC   -!- MISCELLANEOUS: A STRUCTURAL MODEL IS SUGGESTED IN WHICH THE ICE
CC       NUCLEATION PROTEIN DISPLAYS A SYMMETRY RELATED TO THAT OF ICE.
CC   -!- SIMILARITY: BELONGS TO THE BACTERIAL ICE NUCLEATION PROTEIN
CC       FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X52970; CAA37140.1; -.
DR   HSSP; P06620; 1INA.
DR   InterPro; IPR000258; Ice_nucleatn.
DR   Pfam; PF00818; Ice_nucleation; 81.
DR   PRINTS; PR00327; ICENUCLEATN.
DR   PROSITE; PS00314; ICE_NUCLEATION; 57.
KW   Ice nucleation; Repeat; Outer membrane.
SQ   SEQUENCE   1567 AA;  152548 MW;  C8B451D959ECAD63 CRC64;
```

```
  Query Match             6.3%;  Score 173;  DB 1;  Length 1567;
  Best Local Similarity   20.8%;  Pred. No. 0.29;
  Matches   91;  Conservative   67;  Mismatches  220;  Indels   60;  Gaps   15;

Qy        117 TAGTNGDTTVHLNGIGSTLTDMLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKG 176
Db        864 STGTAGADSTLISGYGSTQT-----AGSDSSLTAGYGSTQTARKGSDV----TAGYGSTG 914

Qy        177 VKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGK 236
Db        915 -----TAGADSTLIAGYGSTQTSGSDSSLTAGYGSTQTARKGSDMTAGYGSTGTAGADST 969

Qy        237 LVTGKGKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTK 296
Db        970 LIAGYGSTQTSGS-------------DSSLTAGYGSTQTAREGSDVTAGYGSTGTAGAD 1015

Qy        297 VTFASGNGTTATVSKDDQGNITVKYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISG 356
Db       1016 STLIAGYGSTQTAGSD--SSLTAGY----GSTQTARQGSDVTAGYGSTGTAGADSTLIAG 1069

Qy        357 NVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEG 416
Db       1070 YGSTQTAGSDSSLTAGYGST-QTARQGSDI------TAGYGSTGT-AGADSSLIA--GYG 1119

Qy        417 ALNVGSKDANKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVNGNARAGIAQAIA 476
Db       1120 STQTAGYDSNLTAGYGSTQTAREDSSLTAGYGSTSTAGHDSSLIAGYGSTQTAGYNSILT 1179

Qy        477 TA-GLVQAYLPGKSMMAIGGGTYLGEAGY----AIGYSSISAGG--NWIIKGTASGNSRG 529
Db       1180 TGYGSTQTAQESSSLTAGYGST--STAGYDSTLTAGYGSTQTAGYKSTLTAGYGSNSTAG 1237

Qy        530 H-------FGASASVGYQ 540
Db       1238 HESSLIAGYGSTQIAGYE 1255
```

```
Search completed: October  6, 2003, 09:24:12
Job time : 11.3544 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: October 6, 2003, 09:13:10 ; Search time 39.7345 Seconds
(without alignments)
3513.485 Million cell updates/sec

Title: US-09-771-382-34
Perfect score: 2735
Sequence: 1 TDEDEEEELESVQRSVVGSI.....TASGNSRGHFGASASVGYQW 541

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_23:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2735	100.0	592	2	Q93QY2	Q93qy2 neisseria m
2	2690	98.4	592	2	Q9JPS9	Q9jps9 neisseria m
3	2527	92.4	592	16	Q9JQW4	Q9jqw4 neisseria m
4	2464.5	90.1	594	2	Q9JPI3	Q9jpi3 neisseria m
5	2464.5	90.1	594	2	Q9JPS2	Q9jps2 neisseria m
6	2447.5	89.5	599	2	Q9JPR8	Q9jpr8 neisseria m
7	2426.5	88.7	598	2	Q9JPS0	Q9jps0 neisseria m
8	2426.5	88.7	598	2	Q93QY5	Q93qy5 neisseria m
9	2426.5	88.7	598	2	Q9JPT0	Q9jpt0 neisseria m
10	2423.5	88.6	594	2	Q9JPH7	Q9jph7 neisseria m
11	2412.5	88.2	598	2	Q9JPR9	Q9jpr9 neisseria m
12	2404.5	87.9	594	2	Q93QY4	Q93qy4 neisseria m
13	2401	87.8	590	2	Q9JPS3	Q9jps3 neisseria m
14	2397	87.6	600	2	Q9JPS6	Q9jps6 neisseria m
15	2363.5	86.4	526	2	Q9JPS4	Q9jps4 neisseria m
16	2361.5	86.3	530	2	Q9JPS1	Q9jps1 neisseria m
17	2338	85.5	600	2	Q9JPS5	Q9jps5 neisseria m
18	2330.5	85.2	591	2	Q9JPS7	Q9jps7 neisseria m
19	2330.5	85.2	591	16	Q9JR18	Q9jr18 neisseria m
20	2323.5	85.0	591	2	Q93QY3	Q93qy3 neisseria m
21	2315.5	84.7	592	2	Q9AQF0	Q9aqf0 neisseria m
22	2297.5	84.0	599	2	Q9JPS8	Q9jps8 neisseria m
23	2296.5	84.0	595	2	Q9JPH0	Q9jph0 neisseria m
24	2294.5	83.9	598	2	Q9JPR7	Q9jpr7 neisseria m
25	2256.5	82.5	589	2	Q93QY1	Q93qy1 neisseria m
26	2244.5	82.1	589	2	Q9JPI0	Q9jpi0 neisseria m
27	946	34.6	2353	2	P71401	P71401 haemophilus
28	883	32.3	1204	2	Q8GM76	Q8gm76 haemophilus
29	870.5	31.8	1096	2	Q8GM79	Q8gm79 haemophilus
30	869.5	31.8	1098	2	Q48152	Q48152 haemophilus
31	869.5	31.8	1210	2	Q8GM74	Q8gm74 haemophilus
32	866.5	31.7	1210	2	Q8GM75	Q8gm75 haemophilus
33	588	21.5	1002	2	Q8GM78	Q8gm78 haemophilus
34	585.5	21.4	1004	2	Q8GM77	Q8gm77 haemophilus
35	415	15.2	1299	16	Q9F3X6	Q9f3x6 pasteurella
36	392.5	14.4	2059	16	Q9PD50	Q9pd50 xylella fas
37	378.5	13.8	2314	2	Q8KQM8	Q8kqm8 moraxella c
38	377	13.8	1588	16	Q8XDG4	Q8xdg4 escherichia
39	368.5	13.5	1190	16	Q9PC04	Q9pc04 xylella fas
40	368	13.5	1964	2	Q8KQM9	Q8kqm9 moraxella c
41	367	13.4	1461	16	Q8ZL64	Q8zl64 salmonella
42	363	13.3	1107	16	Q9F2D8	Q9f2d8 salmonella
43	361	13.2	1778	16	Q8FCB2	Q8fcb2 escherichia
44	345	12.6	2712	16	Q9F3X5	Q9f3x5 pasteurella
45	337.5	12.3	688	2	Q8RQ60	Q8rq60 actinobacil

ALIGNMENTS

RESULT 1
Q93QY2
ID Q93QY2 PRELIMINARY; PRT; 592 AA.
AC Q93QY2;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE NhhA outer membrane protein.
GN NHHA.
OS Neisseria meningitidis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=487;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H41;
RA Peak I.R., Srikhanta Y., Dieckelman M., Moxon R., Jennings M.P.;
RT "Identification and characterization of a gene encoding a novel outer
RT membrane protein of Neisseria meningitidis.";
RL Submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF157609; AAK68870.1; -.
DR InterPro; IPR005594; YadA.
DR Pfam; PF03895; YadA; 1.
SQ SEQUENCE 592 AA; 61869 MW; F9403A0B4A18EEA7 CRC64;

Query Match 100.0%; Score 2735; DB 2; Length 592;
Best Local Similarity 100.0%; Pred. No. 4.4e-106;
Matches 541; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 TDEDEEEELESVQRSVVGSIQASMEGSVELETISLSMTNDSKEFVDPYIVVTLKAGDNLK 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         52 TDEDEEEELESVQRSVVGSIQASMEGSVELETISLSMTNDSKEFVDPYIVVTLKAGDNLK 111

Qy         61 IKQNTNENTNASSFTYSLKKDLTGLINVETEKLSFGANGKKVNIISDTKGLNFAKETAGT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        112 IKQNTNENTNASSFTYSLKKDLTGLINVETEKLSFGANGKKVNIISDTKGLNFAKETAGT 171

Qy        121 NGDTTVHLNGIGSTLTDMLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        172 NGDTTVHLNGIGSTLTDMLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPG 231

Qy        181 TTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        232 TTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTG 291

Qy        241 KGKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTKVTFA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        292 KGKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTKVTFA 351

Qy        301 SGNGTTATVSKDDQGNITVKYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSP 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        352 SGNGTTATVSKDDQGNITVKYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSP 411

Qy        361 SKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGALNV 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        412 SKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGALNV 471

Qy        421 GSKDANKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVNGNARAGIAQAIATAGL 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        472 GSKDANKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVNGNARAGIAQAIATAGL 531

Qy        481 VQAYLPGKSMMAIGGGTYLGEAGYAIGYSSISAGGNWIIKGTASGNSRGHFGASASVGYQ 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        532 VQAYLPGKSMMAIGGGTYLGEAGYAIGYSSISAGGNWIIKGTASGNSRGHFGASASVGYQ 591

Qy        541 W 541
              |
Db        592 W 592
```

RESULT 2
Q9JPS9
ID Q9JPS9 PRELIMINARY; PRT; 592 AA.
AC Q9JPS9;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Outer membrane protein GNA992.
GN GNA992.
OS Neisseria meningitidis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=487;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=860800;
RX MEDLINE=20175756; PubMed=10710308;
RA Pizza M., Scarlato V., Masignani V., Giuliani M.M., Arico' B.,
RA Comanducci M., Jennings G.T., Baldi L., Bartolini E., Capecchi B.,
RA Galeotti C.L., Luzzi E., Manetti R., Marchetti E., Mora M., Nuti S.,
RA Ratti G., Santini L., Savino S., Scarselli M., Storni E., Zuo P.,
RA Broeker M., Hundt E., Knapp B., Blair E., Mason T., Tettelin H.,
RA Hood D.W., Jeffries A.C., Saunders N.J., Granoff D.M., Venter C.,
RA Moxon E.R., Grandi G., Rappuoli R.;
RT "Identification of Vaccine Candidates Against Serogroup B
RT Meningococcus by Whole-Genome Sequencing.";
RL Science 287:1816-1820(2000).
DR EMBL; AF226361; AAF42510.1; -.
DR InterPro; IPR005594; YadA.
DR Pfam; PF03895; YadA; 1.
SQ SEQUENCE 592 AA; 61917 MW; 4A3471514FD3C879 CRC64;

Query Match 98.4%; Score 2690; DB 2; Length 592;
Best Local Similarity 98.5%; Pred. No. 3.2e-104;
Matches 533; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

```
Qy          1 TDEDEEEELESVQRSVVGSIQASMEGSVELETISLSMTNDSKEFVDPYIVVTLKAGDNLK 60
              ||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db         52 TDEDEEEELESVQRSVVGSIQASMEGSGELETISLSMTNDSKEFVDPYIVVTLKAGDNLK 111

Qy         61 IKQNTNENTNASSFTYSLKKDLTGLINVETEKLSFGANGKKVNIISDTKGLNFAKETAGT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        112 IKQNTNENTNASSFTYSLKKDLTGLINVETEKLSFGANGKKVNIISDTKGLNFAKETAGT 171

Qy        121 NGDTTVHLNGIGSTLTDMLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        172 NGDTTVHLNGIGSTLTDMLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPG 231

Qy        181 TTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        232 TTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTG 291

Qy        241 KGKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFA 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db        292 KGKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFA 351

Qy        301 SGNGTTATVSKDDQGNITVKYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSP 360
              || |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        352 SGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSP 411

Qy        361 SKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGALNV 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||
Db        412 SKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDKGALNV 471

Qy        421 GSKDANKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVNGNARAGIAQAIATAGL 480
              ||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||
Db        472 GSKDANKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGL 531

Qy        481 VQAYLPGKSMMAIGGGTYLGEAGYAIGYSSISAGGNWIIKGTASGNSRGHFGASASVGYQ 540
              |||||||||||||||| |||||||||||| |||||||||||||||||||||||||||||||
Db        532 VQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQ 591

Qy        541 W 541
              |
Db        592 W 592
```

RESULT 3
Q9JQW4
ID Q9JQW4 PRELIMINARY; PRT; 592 AA.
AC Q9JQW4;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Putative surface fibril protein (Outer membrane protein GNA992).
GN NMA1200 OR GNA992.
OS Neisseria meningitidis (serogroup A), and
OS Neisseria meningitidis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=65699, 487;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Z2491 / Serogroup A / Serotype 4A;
RX MEDLINE=20222556; PubMed=10761919;
RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,
RA Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
RA Davies R.M., Davis P., Devlin K., Feltwell T., Hamlin N., Holroyd S.,
RA Jagels K., Leather S., Moule S., Mungall K., Quail M.A.,
RA Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,
RA Whitehead S., Spratt B.G., Barrell B.G.;
RT "Complete DNA sequence of a serogroup A strain of Neisseria
RT meningitidis Z2491.";
RL Nature 404:502-506(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=205900, BZ133, F6124, AND Z2491;
RX MEDLINE=20175756; PubMed=10710308;
RA Pizza M., Scarlato V., Masignani V., Giuliani M.M., Arico' B.,
RA Comanducci M., Jennings G.T., Baldi L., Bartolini E., Capecchi B.,
RA Galeotti C.L., Luzzi E., Manetti R., Marchetti E., Mora M., Nuti S.,
RA Ratti G., Santini L., Savino S., Scarselli M., Storni E., Zuo P.,
RA Broeker M., Hundt E., Knapp B., Blair E., Mason T., Tettelin H.,
RA Hood D.W., Jeffries A.C., Saunders N.J., Granoff D.M., Venter C.,
RA Moxon E.R., Grandi G., Rappuoli R.;

```
RT   "Identification of Vaccine Candidates Against Serogroup B
RT   Meningococcus by Whole-Genome Sequencing.";
RL   Science 287:1816-1820(2000).
DR   EMBL; AL162755; CAB84461.1; -.
DR   EMBL; AF226357; AAF42506.1; -.
DR   EMBL; AF226365; AAF42514.1; -.
DR   EMBL; AF226373; AAF42522.1; -.
DR   EMBL; AF226386; AAF42535.1; -.
DR   InterPro; IPR005594; YadA.
DR   Pfam; PF03895; YadA; 1.
KW   Complete proteome.
SQ   SEQUENCE   592 AA;  61745 MW;  AD22E2F5EEF8F754 CRC64;

  Query Match             92.4%;  Score 2527;  DB 16;  Length 592;
  Best Local Similarity   93.4%;  Pred. No. 1.8e-97;
  Matches  507;  Conservative    9;  Mismatches   23;  Indels    4;  Gaps    2;

Qy          1 TDEDEEEELESVQRSVVGSIQASMEGSVELETISLSMTNDSKEFVDPYIVVTLKAGDNLK 60
              ||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db         52 TDEDEEEELESVQRSVVGSIQASMEGSGELETISLSMTNDSKEFVDPYIVVTLKAGDNLK 111

Qy         61 IKQNTNENTNASSFTYSLKKDLTGLINVETEKLSFGANGKKVNIISDTKGLNFAKETAGT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        112 IKQNTNENTNASSFTYSLKKDLTGLINVETEKLSFGANGKKVNIISDTKGLNFAKETAGT 171

Qy        121 NGDTTVHLNGIGSTLTDMLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPG 180
              |||||||||||||||| |  : |:     :  |     ||||:||||||||||||| |
Db        172 NGDTTVHLNGIGSTLTDTLAGSSASHVDAGNQST--HYTRAASIKDVLNAGWNIKGVKTG 229

Qy        181 TTA--SDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLV 238
              :|   |:||||||||||||||||||||||||||||:||||||||||||||||||||||||
Db        230 STTGQSENVDFVRTYDTVEFLSADTKTTTVNVESKDNGKRTEVKIGAKTSVIKEKDGKLV 289

Qy        239 TGKGKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTKVT 298
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db        290 TGKGKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVT 349

Qy        299 FASGNGTTATVSKDDQGNITVKYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNV 358
              |||| |||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        350 FASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNV 409

Qy        359 SPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGAL 418
              |||||||||||||||||||:|||||||||||| ||||||||||||||||||||||||||||
Db        410 SPSKGKMDETVNINAGNNIEISRNGKNIDIATSMAPQFSSVSLGAGADAPTLSVDDEGAL 469

Qy        419 NVGSKDANKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVNGNARAGIAQAIATA 478
              ||||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||
Db        470 NVGSKDANKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATA 529

Qy        479 GLVQAYLPGKSMMAIGGGTYLGEAGYAIGYSSISAGGNWIIKGTASGNSRGHFGASASVG 538
              |||||||||||||||||| |||||||||||| |||||||||||||||||||||||||||||
Db        530 GLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVG 589

Qy        539 YQW 541
              |||
Db        590 YQW 592


RESULT 4
Q9JPI3
ID   Q9JPI3      PRELIMINARY;      PRT;   594 AA.
AC   Q9JPI3;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Outer membrane protein GNA992.
GN   GNA992.
OS   Neisseria meningitidis.
OC   Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC   Neisseriaceae; Neisseria.
OX   NCBI_TaxID=487;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NG3/88, and BZ232;
RX   MEDLINE=20175756; PubMed=10710308;
RA   Pizza M., Scarlato V., Masignani V., Giuliani M.M., Arico' B.,
RA   Comanducci M., Jennings G.T., Baldi L., Bartolini E., Capecchi B.,
RA   Galeotti C.L., Luzzi E., Manetti R., Marchetti E., Mora M., Nuti S.,
RA   Ratti G., Santini L., Savino S., Scarselli M., Storni E., Zuo P.,
RA   Broeker M., Hundt E., Knapp B., Blair E., Mason T., Tettelin H.,
RA   Hood D.W., Jeffries A.C., Saunders N.J., Granoff D.M., Venter C.,
RA   Moxon E.R., Grandi G., Rappuoli R.;
RT   "Identification of Vaccine Candidates Against Serogroup B
RT   Meningococcus by Whole-Genome Sequencing.";
RL   Science 287:1816-1820(2000).
DR   EMBL; AF226376; AAF42525.1; -.
DR   EMBL; AF226369; AAF42518.1; -.
DR   InterPro; IPR005594; YadA.
DR   Pfam; PF03895; YadA; 1.
SQ   SEQUENCE   594 AA;  62086 MW;  1B25E03B90D04B46 CRC64;

  Query Match             90.1%;  Score 2464.5;  DB 2;  Length 594;
  Best Local Similarity   90.7%;  Pred. No. 7.2e-95;
  Matches  498;  Conservative   11;  Mismatches   27;  Indels   13;  Gaps    3;

Qy          1 TDEDEEEELESVQRSV-VGSIQASMEGSVELETISLSMTNDS-------KEFVDPYIVVT 52
              | :|::  || |||:  | |  |  ||: | |     :| ||       |: |     :|
Db         51 TTDDDDLYLEPVQRTAPVLSFHADSEGTGEKE-----VTEDSNWGVYFDKKGVLTAGTIT 105

Qy         53 LKAGDNLKIKQNTNENTNASSFTYSLKKDLTGLINVETEKLSFGANGKKVNIISDTKGLN 112
              |||||||||||||:|||||||||||||||| | :|||||||||||||||| |||||||
Db        106 LKAGDNLKIKQNTDENTNASSFTYSLKKDLTDLTSVETEKLSFGANGKKVNITSDTKGLN 165

Qy        113 FAKETAGTNGDTTVHLNGIGSTLTDMLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGW 172
              |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db        166 FAKETAGTNGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGW 225

Qy        173 NIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKE 232
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        226 NIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKE 285

Qy        233 KDGKLVTGKGKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVT 292
              ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        286 KDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVT 345

Qy        293 SGTKVTFASGNGTTATVSKDDQGNITVKYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGK 352
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        346 SGTKVTFASGNGTTATVSKDDQGNITVKYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGK 405

Qy        353 VISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSV 412
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        406 VISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSV 465

Qy        413 DDEGALNVGSKDANKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVNGNARAGIA 472
              |||||||||||||||||||||||||||||||||||||||||||||| ||||:||||||||
Db        466 DDEGALNVGSKDANKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNHIDNVDGNARAGIA 525

Qy        473 QAIATAGLVQAYLPGKSMMAIGGGTYLGEAGYAIGYSSISAGGNWIIKGTASGNSRGHFG 532
              |||||||||||||||||||||||| |||||||||||| |||||||||||||||||||||||
Db        526 QAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFG 585

Qy        533 ASASVGYQW 541
              |||||||||
Db        586 ASASVGYQW 594


RESULT 5
Q9JPS2
ID   Q9JPS2      PRELIMINARY;      PRT;   594 AA.
AC   Q9JPS2;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Outer membrane protein GNA992.
GN   GNA992.
OS   Neisseria meningitidis.
```

```
OC   Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC   Neisseriaceae; Neisseria.
OX   NCBI_TaxID=487;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NGE31;
RX   MEDLINE=20175756; PubMed=10710308;
RA   Pizza M., Scarlato V., Masignani V., Giuliani M.M., Arico' B.,
RA   Comanducci M., Jennings G.T., Baldi L., Bartolini E., Capecchi B.,
RA   Galeotti C.L., Luzzi E., Manetti R., Marchetti E., Mora M., Nuti S.,
RA   Ratti G., Santini L., Savino S., Scarselli M., Storni E., Zuo P.,
RA   Broeker M., Hundt E., Knapp B., Blair E., Mason T., Tettelin H.,
RA   Hood D.W., Jeffries A.C., Saunders N.J., Granoff D.M., Venter C.,
RA   Moxon E.R., Grandi G., Rappuoli R.;
RT   "Identification of Vaccine Candidates Against Serogroup B
RT   Meningococcus by Whole-Genome Sequencing.";
RL   Science 287:1816-1820(2000).
DR   EMBL; AF226379; AAF42528.1; -.
DR   InterPro; IPR005594; YadA.
DR   Pfam; PF03895; YadA; 1.
SQ   SEQUENCE   594 AA;  62114 MW;  1E2A63A78F53D256 CRC64;

  Query Match             90.1%;  Score 2464.5;  DB 2;  Length 594;
  Best Local Similarity   90.7%;  Pred. No. 7.2e-95;
  Matches  498;  Conservative   11;  Mismatches   27;  Indels   13;  Gaps    3;

Qy          1 TDEDEEEELESVQRSV-VGSIQASMEGSVELETISLSMTNDS-------KEFVDPYIVVT 52
              | :|::  || |||:  | |  |  ||: | |     :| ||       |: |     :|
Db         51 TTDDDDLYLEPVQRTAPVLSFHADSEGTGEKE-----VTEDSNWGVYFDKKGVLTAGTIT 105

Qy         53 LKAGDNLKIKQNTNENTNASSFTYSLKKDLTGLINVETEKLSFGANGKKVNIISDTKGLN 112
              ||||||||||||:|||||||||||||||| | :|||||||||||||||| |||||||
Db        106 LKAGDNLKIKQNTDENTNASSFTYSLKKDLTDLTSVETEKLSFGANGKKVNITSDTKGLN 165

Qy        113 FAKETAGTNGDTTVHLNGIGSTLTDMLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGW 172
              |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db        166 FAKETAGTNGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGW 225

Qy        173 NIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKE 232
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        226 NIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKE 285

Qy        233 KDGKLVTGKGKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVT 292
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        286 KDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVT 345

Qy        293 SGTKVTFASGNGTTATVSKDDQGNITVKYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGK 352
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        346 SGTKVTFASGNGTTATVSKDDQGNITVKYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGK 405

Qy        353 VISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSV 412
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        406 VISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSV 465

Qy        413 DDEGALNVGSKDANKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVNGNARAGIA 472
              |||||||||||||||||||||||||||||||||||||||||||| ||||:||||||||
Db        466 DDEGALNVGSKDANKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNHIDNVDGNARAGIA 525

Qy        473 QAIATAGLVQAYLPGKSMMAIGGGTYLGEAGYAIGYSSISAGGNWIIKGTASGNSRGHFG 532
              |||||||||||||||||||||| |||||||||||| ||||||||||||||||||||||
Db        526 QAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFG 585

Qy        533 ASASVGYQW 541
              |||||||||
Db        586 ASASVGYQW 594


RESULT 6
Q9JPR8
ID   Q9JPR8      PRELIMINARY;      PRT;   599 AA.
AC   Q9JPR8;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Outer membrane protein GNA992 (NhhA outer membrane protein).
GN   GNA992 OR NHHA.
OS   Neisseria meningitidis.
OC   Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC   Neisseriaceae; Neisseria.
OX   NCBI_TaxID=487;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NGH38;
RX   MEDLINE=20175756; PubMed=10710308;
RA   Pizza M., Scarlato V., Masignani V., Giuliani M.M., Arico' B.,
RA   Comanducci M., Jennings G.T., Baldi L., Bartolini E., Capecchi B.,
RA   Galeotti C.L., Luzzi E., Manetti R., Marchetti E., Mora M., Nuti S.,
RA   Ratti G., Santini L., Savino S., Scarselli M., Storni E., Zuo P.,
RA   Broeker M., Hundt E., Knapp B., Blair E., Mason T., Tettelin H.,
RA   Hood D.W., Jeffries A.C., Saunders N.J., Granoff D.M., Venter C.,
RA   Moxon E.R., Grandi G., Rappuoli R.;
RT   "Identification of Vaccine Candidates Against Serogroup B
RT   Meningococcus by Whole-Genome Sequencing.";
RL   Science 287:1816-1820(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=H38;
RA   Peak I.R., Srikhanta Y., Dieckelman M., Moxon R., Jennings M.P.;
RT   "Identification and characterization of a gene encoding a novel outer
RT   membrane protein of Neisseria meningitidis.";
RL   Submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF226383; AAF42532.1; -.
DR   EMBL; AF157608; AAK68869.1; -.
DR   InterPro; IPR005594; YadA.
DR   Pfam; PF03895; YadA; 1.
SQ   SEQUENCE   599 AA;  62844 MW;  BBA16EBF53C1970C CRC64;

  Query Match             89.5%;  Score 2447.5;  DB 2;  Length 599;
  Best Local Similarity   90.3%;  Pred. No. 3.7e-94;
  Matches  496;  Conservative   11;  Mismatches   33;  Indels    9;  Gaps    4;

Qy          1 TDEDEEEELESVQRS-VVGSIQASMEGSVELET---ISLSMTNDSKEFVDPYIVVTLKAG 56
              |||||||||| | || :|       ||: | |:   |  |:  |:   :     ||||||
Db         52 TDEDEEEELEPVVRSALVLQFMIDKEGNGENESTGNIGWSIYYDNHNTLHG-ATVTLKAG 110

Qy         57 DNLKIKQ----NTNENTNASSFTYSLKKDLTGLINVETEKLSFGANGKKVNIISDTKGLN 112
              |||||||    ||||||| |||||||||||| | :||||||||||| |||| |||||||
Db        111 DNLKIKQNTNKNTNENTNDSSFTYSLKKDLTDLTSVETEKLSFGANGNKVNITSDTKGLN 170

Qy        113 FAKETAGTNGDTTVHLNGIGSTLTDMLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGW 172
              |||||||||||||||||||||||| ||||||||||||||||||:|||||||||||||||
Db        171 FAKETAGTNGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDKKKRAASVKDVLNAGW 230

Qy        173 NIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKE 232
              |||||||||||||||||||||||| ||||||||||||||||||:|||||||||||||||
Db        231 NIKGVKPGTTASDNVDFVHTYDTVEFLSADTKTTTVNVESKDNGKRTEVKIGAKTSVIKE 290

Qy        233 KDGKLVTGKGKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVT 292
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        291 KDGKLVTGKGKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVT 350

Qy        293 SGTKVTFASGNGTTATVSKDDQGNITVKYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGK 352
              |||  |||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db        351 SGTNVTFASGKGTTATVSKDDQGNITVKYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGK 410

Qy        353 VISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSV 412
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        411 VISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSV 470

Qy        413 DDEGALNVGSKDANKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVNGNARAGIA 472
              ||:|||||||||||||||||||||||||||||||||||||||||||||:||||||||
Db        471 DDKGALNVGSKDANKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIA 530

Qy        473 QAIATAGLVQAYLPGKSMMAIGGGTYLGEAGYAIGYSSISAGGNWIIKGTASGNSRGHFG 532
              |||||||||||||||||||||| |||||||||||| ||||||||||||||||||||||
Db        531 QAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFG 590
```

```
Qy        533 ASASVGYQW 541
              |||||||||
Db        591 ASASVGYQW 599


RESULT 7
Q9JPS0
ID   Q9JPS0      PRELIMINARY;      PRT;   598 AA.
AC   Q9JPS0;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Outer membrane protein GNA992 (NhhA outer membrane protein).
GN   GNA992 OR NHHA.
OS   Neisseria meningitidis.
OC   Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC   Neisseriaceae; Neisseria.
OX   NCBI_TaxID=487;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NGH15;
RX   MEDLINE=20175756; PubMed=10710308;
RA   Pizza M., Scarlato V., Masignani V., Giuliani M.M., Arico' B.,
RA   Comanducci M., Jennings G.T., Baldi L., Bartolini E., Capecchi B.,
RA   Galeotti C.L., Luzzi E., Manetti R., Marchetti E., Mora M., Nuti S.,
RA   Ratti G., Santini L., Savino S., Scarselli M., Storni E., Zuo P.,
RA   Broeker M., Hundt E., Knapp B., Blair E., Mason T., Tettelin H.,
RA   Hood D.W., Jeffries A.C., Saunders N.J., Granoff D.M., Venter C.,
RA   Moxon E.R., Grandi G., Rappuoli R.;
RT   "Identification of Vaccine Candidates Against Serogroup B
RT   Meningococcus by Whole-Genome Sequencing.";
RL   Science 287:1816-1820(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=H15;
RA   Peak I.R., Srikhanta Y., Dieckelman M., Moxon R., Jennings M.P.;
RT   "Identification and characterization of a gene encoding a novel outer
RT   membrane protein of Neisseria meningitidis.";
RL   Submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF226381; AAF42530.1; -.
DR   EMBL; AF157607; AAK68868.1; -.
DR   InterPro; IPR005594; YadA.
DR   Pfam; PF03895; YadA; 1.
SQ   SEQUENCE   598 AA;  62763 MW;  E6C7AEF0BB8A63CB CRC64;

  Query Match             88.7%;  Score 2426.5;  DB 2;  Length 598;
  Best Local Similarity   89.2%;  Pred. No. 2.7e-93;
  Matches  493;  Conservative   11;  Mismatches   28;  Indels   21;  Gaps    5;

Qy          3 EDEEEELESVQR-SVVGSIQASMEGSVELETISLSMTNDSKEFVDPYI---------VVT 52
              :|::  || ||| :|| | ::  ||: | |      | ||   |  |           :|
Db         53 DDDDLYLEPVQRTAVVLSFRSDKEGTGEKEG-----TEDSNWAV--YFDEKRVLKAGAIT 105

Qy         53 LKAGDNLKIKQNTNENTNA----SSFTYSLKKDLTGLINVETEKLSFGANGKKVNIISDT 108
              ||||||||||||||||||     |||||||||||| | :|||||||||||| |||| |||
Db        106 LKAGDNLKIKQNTNENTNENTNDSSFTYSLKKDLTDLTSVETEKLSFGANGNKVNITSDT 165

Qy        109 KGLNFAKETAGTNGDTTVHLNGIGSTLTDMLLNTGATTNVTNDNVTDDEKKRAASVKDVL 168
              |||||||||||||| |||||||||||| ||||||||||||||||||||||||||||||
Db        166 KGLNFAKETAGTNGDPTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVL 225

Qy        169 NAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTS 228
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        226 NAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTS 285

Qy        229 VIKEKDGKLVTGKGKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKF 288
              ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        286 VIKEKDGKLVTGKGKDENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKF 345

Qy        289 ETVTSGTKVTFASGNGTTATVSKDDQGNITVKYDVNVGDALNVNQLQNSGWNLDSKAVAG 348
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        346 ETVTSGTKVTFASGNGTTATVSKDDQGNITVKYDVNVGDALNVNQLQNSGWNLDSKAVAG 405

Qy        349 SSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAP 408
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        406 SSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAP 465

Qy        409 TLSVDDEGALNVGSKDANKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVNGNAR 468
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||:||||
Db        466 TLSVDDEGALNVGSKDANKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNAR 525

Qy        469 AGIAQAIATAGLVQAYLPGKSMMAIGGGTYLGEAGYAIGYSSISAGGNWIIKGTASGNSR 528
              |||||||||||| |||||||||||||||| |||||||||||  |||:||||||||||
Db        526 AGIAQAIATAGLAQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDTGNWVIKGTASGNSR 585

Qy        529 GHFGASASVGYQW 541
              |||||||||||||
Db        586 GHFGASASVGYQW 598


RESULT 8
Q93QY5
ID   Q93QY5      PRELIMINARY;      PRT;   598 AA.
AC   Q93QY5;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   NhhA outer membrane protein.
GN   NHHA.
OS   Neisseria meningitidis.
OC   Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC   Neisseriaceae; Neisseria.
OX   NCBI_TaxID=487;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BZ10;
RA   Peak I.R., Srikhanta Y., Dieckelman M., Moxon R., Jennings M.P.;
RT   "Identification and characterization of a gene encoding a novel outer
RT   membrane protein of Neisseria meningitidis.";
RL   Submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF157603; AAK68864.1; -.
DR   InterPro; IPR005594; YadA.
DR   Pfam; PF03895; YadA; 1.
SQ   SEQUENCE   598 AA;  62687 MW;  18CEFFE6410A15DF CRC64;

  Query Match             88.7%;  Score 2426.5;  DB 2;  Length 598;
  Best Local Similarity   89.0%;  Pred. No. 2.7e-93;
  Matches  492;  Conservative   12;  Mismatches   28;  Indels   21;  Gaps    5;

Qy          3 EDEEEELESVQR-SVVGSIQASMEGSVELETISLSMTNDSKEFVDPYI---------VVT 52
              :|::  || ||| :|| | ::  ||: | |      | ||   |  |           :|
Db         53 DDDDLYLEPVQRTAVVLSFRSDKEGTGEKEG-----TEDSNWAV--YFDEKRVLKAGAIT 105

Qy         53 LKAGDNLKIKQNTNENTNA----SSFTYSLKKDLTGLINVETEKLSFGANGKKVNIISDT 108
              ||||||||||||||||||     |||||||||||| | :|||||||||||| |||| |||
Db        106 LKAGDNLKIKQNTNENTNENTNDSSFTYSLKKDLTDLTSVETEKLSFGANGNKVNITSDT 165

Qy        109 KGLNFAKETAGTNGDTTVHLNGIGSTLTDMLLNTGATTNVTNDNVTDDEKKRAASVKDVL 168
              |||||||||||||| |||||||||||| ||||||||||||||||||||||||||||||
Db        166 KGLNFAKETAGTNGDPTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVL 225

Qy        169 NAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTS 228
              ||||||||||||||||||||||||||||||||||||||||||||||||||||:||||||||
Db        226 NAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKRTEVKIGAKTS 285

Qy        229 VIKEKDGKLVTGKGKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKF 288
              ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        286 VIKEKDGKLVTGKGKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKF 345

Qy        289 ETVTSGTKVTFASGNGTTATVSKDDQGNITVKYDVNVGDALNVNQLQNSGWNLDSKAVAG 348
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        346 ETVTSGTKVTFASGNGTTATVSKDDQGNITVKYDVNVGDALNVNQLQNSGWNLDSKAVAG 405

Qy        349 SSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAP 408
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        406 SSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAP 465

Qy        409 TLSVDDEGALNVGSKDANKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVNGNAR 468
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||:||||
Db        466 TLSVDDEGALNVGSKDANKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNAR 525

Qy        469 AGIAQAIATAGLVQAYLPGKSMMAIGGGTYLGEAGYAIGYSSISAGGNWIIKGTASGNSR 528
              |||||||||||| |||||||||||||||| |||||||||||||  |||:||||||||||
Db        526 AGIAQAIATAGLAQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDTGNWVIKGTASGNSR 585

Qy        529 GHFGASASVGYQW 541
              |||| ||||||||
Db        586 GHFGTSASVGYQW 598


RESULT 9
Q9JPT0
ID   Q9JPT0      PRELIMINARY;      PRT;   598 AA.
AC   Q9JPT0;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Outer membrane protein GNA992.
GN   GNA992.
OS   Neisseria meningitidis.
OC   Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC   Neisseriaceae; Neisseria.
OX   NCBI_TaxID=487;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=2996;
RX   MEDLINE=20175756; PubMed=10710308;
RA   Pizza M., Scarlato V., Masignani V., Giuliani M.M., Arico' B.,
RA   Comanducci M., Jennings G.T., Baldi L., Bartolini E., Capecchi B.,
RA   Galeotti C.L., Luzzi E., Manetti R., Marchetti E., Mora M., Nuti S.,
RA   Ratti G., Santini L., Savino S., Scarselli M., Storni E., Zuo P.,
RA   Broeker M., Hundt E., Knapp B., Blair E., Mason T., Tettelin H.,
RA   Hood D.W., Jeffries A.C., Saunders N.J., Granoff D.M., Venter C.,
RA   Moxon E.R., Grandi G., Rappuoli R.;
RT   "Identification of Vaccine Candidates Against Serogroup B
RT   Meningococcus by Whole-Genome Sequencing.";
RL   Science 287:1816-1820(2000).
DR   EMBL; AF226359; AAF42508.1; -.
DR   InterPro; IPR005594; YadA.
DR   Pfam; PF03895; YadA; 1.
SQ   SEQUENCE   598 AA;  62763 MW;  63A6A3BD7F0F2EE3 CRC64;

  Query Match           88.7%;  Score 2426.5;  DB 2;  Length 598;
  Best Local Similarity   89.0%;  Pred. No. 2.7e-93;
  Matches  492;  Conservative   12;  Mismatches   28;  Indels   21;  Gaps    5;

Qy          3 EDEEEELESVQR-SVVGSIQASMEGSVELETISLSMTNDSKEFVDPYI---------VVT 52
              :|::  || ||| :|| | ::  ||: | |      | ||   |  |          :|
Db         53 DDDDLYLEPVQRTAVVLSFRSDKEGTGEKEG-----TEDSNWAV--YFDEKRVLKAGAIT 105

Qy         53 LKAGDNLKIKQNTNENTNA----SSFTYSLKKDLTGLINVETEKLSFGANGKKVNIISDT 108
              ||||||||||||||||||     |||||||||||| | :||||||||||||| |||| |||
Db        106 LKAGDNLKIKQNTNENTNENTNDSSFTYSLKKDLTDLTSVETEKLSFGANGNKVNITSDT 165

Qy        109 KGLNFAKETAGTNGDTTVHLNGIGSTLTDMLLNTGATTNVTNDNVTDDEKKRAASVKDVL 168
              |||||||||||||| ||||||||||||| ||||||||||||||||||||||||||||||
Db        166 KGLNFAKETAGTNGDPTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVL 225

Qy        169 NAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTS 228
              ||||||||||||||||||||||||||||||||||||||||||||:||||||||||||||
Db        226 NAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKRTEVKIGAKTS 285

Qy        229 VIKEKDGKLVTGKGKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKF 288
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        286 VIKEKDGKLVTGKGKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKF 345

Qy        289 ETVTSGTKVTFASGNGTTATVSKDDQGNITVKYDVNVGDALNVNQLQNSGWNLDSKAVAG 348
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        346 ETVTSGTKVTFASGNGTTATVSKDDQGNITVKYDVNVGDALNVNQLQNSGWNLDSKAVAG 405

Qy        349 SSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAP 408
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        406 SSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAP 465

Qy        409 TLSVDDEGALNVGSKDANKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVNGNAR 468
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||:||||
Db        466 TLSVDDEGALNVGSKDANKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNAR 525

Qy        469 AGIAQAIATAGLVQAYLPGKSMMAIGGGTYLGEAGYAIGYSSISAGGNWIIKGTASGNSR 528
              |||||||||||| |||||||||||||||| |||||||||||||  |||:||||||||||
Db        526 AGIAQAIATAGLAQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDTGNWVIKGTASGNSR 585

Qy        529 GHFGASASVGYQW 541
              |||| ||||||||
Db        586 GHFGTSASVGYQW 598


RESULT 10
Q9JPH7
ID   Q9JPH7      PRELIMINARY;      PRT;   594 AA.
AC   Q9JPH7;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Outer membrane protein GNA992 (NhhA outer membrane protein).
GN   GNA992 OR NHHA.
OS   Neisseria meningitidis.
OC   Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC   Neisseriaceae; Neisseria.
OX   NCBI_TaxID=487;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BZ198, and 297-0;
RX   MEDLINE=20175756; PubMed=10710308;
RA   Pizza M., Scarlato V., Masignani V., Giuliani M.M., Arico' B.,
RA   Comanducci M., Jennings G.T., Baldi L., Bartolini E., Capecchi B.,
RA   Galeotti C.L., Luzzi E., Manetti R., Marchetti E., Mora M., Nuti S.,
RA   Ratti G., Santini L., Savino S., Scarselli M., Storni E., Zuo P.,
RA   Broeker M., Hundt E., Knapp B., Blair E., Mason T., Tettelin H.,
RA   Hood D.W., Jeffries A.C., Saunders N.J., Granoff D.M., Venter C.,
RA   Moxon E.R., Grandi G., Rappuoli R.;
RT   "Identification of Vaccine Candidates Against Serogroup B
RT   Meningococcus by Whole-Genome Sequencing.";
RL   Science 287:1816-1820(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BZ198;
RA   Peak I.R., Srikhanta Y., Dieckelman M., Moxon R., Jennings M.P.;
RT   "Identification and characterization of a gene encoding a novel outer
RT   membrane protein of Neisseria meningitidis.";
RL   Submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF226368; AAF42517.1; -.
DR   EMBL; AF226358; AAF42507.1; -.
DR   EMBL; AF157604; AAK68865.1; -.
DR   InterPro; IPR005594; YadA.
DR   Pfam; PF03895; YadA; 1.
SQ   SEQUENCE   594 AA;  62361 MW;  436BDDED68263C5C CRC64;

  Query Match           88.6%;  Score 2423.5;  DB 2;  Length 594;
  Best Local Similarity   89.4%;  Pred. No. 3.6e-93;
  Matches  491;  Conservative   10;  Mismatches   31;  Indels   17;  Gaps    4;

Qy          3 EDEEEELESVQR-SVVGSIQASMEGSVELETISLSMTNDSKEFVDPYI---------VVT 52
              :|::  || ||| :|| | ::  ||: | |      | ||   |  |          :|
Db         53 DDDDLYLEPVQRTAVVLSFRSDKEGTGEKEG-----TEDSNWAV--YFDEKRVLKAGAIT 105

Qy         53 LKAGDNLKIKQNTNENTNASSFTYSLKKDLTGLINVETEKLSFGANGKKVNIISDTKGLN 112
              |||||||||||||||||| |||||||||||| | :|||||||||||| |||| |||||||
Db        106 LKAGDNLKIKQNTNENTNDSSFTYSLKKDLTDLTSVETEKLSFGANGNKVNITSDTKGLN 165

Qy        113 FAKETAGTNGDTTVHLNGIGSTLTDMLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGW 172
```

```
Db        166 FAKETAGTNGDPTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGW 225

Qy        173 NIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKE 232
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        226 NIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKE 285

Qy        233 KDGKLVTGKGKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVT 292
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        286 KDGKLVTGKGKDENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVT 345

Qy        293 SGTKVTFASGNGTTATVSKDDQGNITVKYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGK 352
              ||| |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        346 SGTNVTFASGKGTTATVSKDDQGNITVKYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGK 405

Qy        353 VISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSV 412
              ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        406 VISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMAPQFSSVSLGAGADAPTLSV 465

Qy        413 DDEGALNVGSKDANKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVNGNARAGIA 472
              |||||||||||| ||||||||||||||||||||||||||||||||:|||||||
Db        466 DDEGALNVGSKDTNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIA 525

Qy        473 QAIATAGLVQAYLPGKSMMAIGGGTYLGEAGYAIGYSSISAGGNWIIKGTASGNSRGHFG 532
              |||||||||||||||||||| || |||||||||||| ||||||||||||||||||||||||
Db        526 QAIATAGLVQAYLPGKSMMAIGGDTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFG 585

Qy        533 ASASVGYQW 541
              |||||||||
Db        586 ASASVGYQW 594
```

RESULT 11
Q9JPR9
ID Q9JPR9 PRELIMINARY; PRT; 598 AA.
AC Q9JPR9;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Outer membrane protein GNA992.
GN GNA992.
OS Neisseria meningitidis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=487;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NGH36;
RX MEDLINE=20175756; PubMed=10710308;
RA Pizza M., Scarlato V., Masignani V., Giuliani M.M., Arico' B.,
RA Comanducci M., Jennings G.T., Baldi L., Bartolini E., Capecchi B.,
RA Galeotti C.L., Luzzi E., Manetti R., Marchetti E., Mora M., Nuti S.,
RA Ratti G., Santini L., Savino S., Scarselli M., Storni E., Zuo P.,
RA Broeker M., Hundt E., Knapp B., Blair E., Mason T., Tettelin H.,
RA Hood D.W., Jeffries A.C., Saunders N.J., Granoff D.M., Venter C.,
RA Moxon E.R., Grandi G., Rappuoli R.;
RT "Identification of Vaccine Candidates Against Serogroup B
RT Meningococcus by Whole-Genome Sequencing.";
RL Science 287:1816-1820(2000).
DR EMBL; AF226382; AAF42531.1; -.
DR InterPro; IPR005594; YadA.
DR Pfam; PF03895; YadA; 1.
SQ SEQUENCE 598 AA; 62718 MW; 9095F8E31AD7C76D CRC64;

Query Match 88.2%; Score 2412.5; DB 2; Length 598;
Best Local Similarity 88.8%; Pred. No. 1e-92;
Matches 491; Conservative 10; Mismatches 31; Indels 21; Gaps 5;

```
Qy          3 EDEEEELESVQR-SVVGSIQASMEGSVELETISLSMTNDSKEFVDPYI---------VVT 52
              :|::  || ||| :|| | ::  ||: | |      | ||   |  |           :|
Db         53 DDDDLYLEPVQRTAVVLSFRSDKEGTGEKEG-----TEDSNWAV--YFDEKRVLKAGAIT 105

Qy         53 LKAGDNLKIKQNTNENTNA----SSFTYSLKKDLTGLINVETEKLSFGANGKKVNIISDT 108
```

```
Db        106 LKAGDNLKIKQNTNENTNENTNDSSFTYSLKKDLTDLTSVETEKLSFGANGNKVNITSDT 165

Qy        109 KGLNFAKETAGTNGDTTVHLNGIGSTLTDMLLNTGATTNVTNDNVTDDEKKRAASVKDVL 168
              |||||||||||||| |||||||||||| ||||||||||||||||||||||||||||||||
Db        166 KGLNFAKETAGTNGDPTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVL 225

Qy        169 NAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTS 228
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        226 NAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTS 285

Qy        229 VIKEKDGKLVTGKGKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKF 288
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        286 VIKEKDGKLVTGKGKDENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKF 345

Qy        289 ETVTSGTKVTFASGNGTTATVSKDDQGNITVKYDVNVGDALNVNQLQNSGWNLDSKAVAG 348
              ||||||| |||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        346 ETVTSGTNVTFASGKGTTATVSKDDQGNITVKYDVNVGDALNVNQLQNSGWNLDSKAVAG 405

Qy        349 SSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAP 408
              |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        406 SSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMAPQFSSVSLGAGADAP 465

Qy        409 TLSVDDEGALNVGSKDANKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVNGNAR 468
              |||||||||||||| ||||||||||||||||||||||||||||||||||:|||
Db        466 TLSVDDEGALNVGSKDTNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNAL 525

Qy        469 AGIAQAIATAGLVQAYLPGKSMMAIGGGTYLGEAGYAIGYSSISAGGNWIIKGTASGNSR 528
              |||||||||||||||||||||||||||||||||||| |||||||||||| ||||||||||
Db        526 AGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSR 585

Qy        529 GHFGASASVGYQW 541
              |||||||||||||
Db        586 GHFGASASVGYQW 598
```

RESULT 12
Q93QY4
ID Q93QY4 PRELIMINARY; PRT; 594 AA.
AC Q93QY4;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE NhhA outer membrane protein.
GN NHHA.
OS Neisseria meningitidis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=487;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EG327;
RA Peak I.R., Srikhanta Y., Dieckelman M., Moxon R., Jennings M.P.;
RT "Identification and characterization of a gene encoding a novel outer
RT membrane protein of Neisseria meningitidis.";
RL Submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF157605; AAK68866.1; -.
DR InterPro; IPR005594; YadA.
DR Pfam; PF03895; YadA; 1.
SQ SEQUENCE 594 AA; 62297 MW; 9DDD48B04B3A8EA2 CRC64;

Query Match 87.9%; Score 2404.5; DB 2; Length 594;
Best Local Similarity 88.7%; Pred. No. 2.2e-92;
Matches 487; Conservative 15; Mismatches 34; Indels 13; Gaps 3;

```
Qy          1 TDEDEEEELESVQR-SVVGSIQASMEGSVELETISLSMTNDS-------KEFVDPYIVVT 52
              | :|::  || ||| :|| | ::  ||: | |     :| ||       |: |     :|
Db         51 TTDDDDLYLEPVQRTAVVLSFRSDKEGTGEKE-----VTEDSNWGVYFDKKGVLTAGTIT 105

Qy         53 LKAGDNLKIKQNTNENTNASSFTYSLKKDLTGLINVETEKLSFGANGKKVNIISDTKGLN 112
              |||||||||||||||||||||||||||||| | :| |||||| ||  |||| |||||||
Db        106 LKAGDNLKIKQNTNENTNASSFTYSLKKDLTDLTSVGTEKLSFSANSNKVNITSDTKGLN 165
```

```
Qy        113 FAKETAGTNGDTTVHLNGIGSTLTDMLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGW 172
              |||:|| |||||||||||||||||| |||||||||||||||||||||||||||||||||
Db        166 FAKKTAETNGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGW 225

Qy        173 NIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKE 232
              ||||||||||||||||||||||||||||||||||||||||||:|||||||||||||||||
Db        226 NIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKRTEVKIGAKTSVIKE 285

Qy        233 KDGKLVTGKGKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVT 292
              |||||||| |||| ||||:|||||||||||||||||||||||||||||||||||||||||
Db        286 KDGKLVTGKDKGENDSSTDKGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVT 345

Qy        293 SGTKVTFASGNGTTATVSKDDQGNITVKYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGK 352
              ||| |||||| ||||||||||||||| ||||||||||||||||||||||||||||||||
Db        346 SGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGK 405

Qy        353 VISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSV 412
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        406 VISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSV 465

Qy        413 DDEGALNVGSKDANKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVNGNARAGIA 472
              |||||||||||||||||||||||||||||||||||||||||||| ||||:||||||||
Db        466 DDEGALNVGSKDANKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNHIDNVDGNARAGIA 525

Qy        473 QAIATAGLVQAYLPGKSMMAIGGGTYLGEAGYAIGYSSISAGGNWIIKGTASGNSRGHFG 532
              |||||||||||||||||||||||| |||||||||||| |||||||||||||||||||||
Db        526 QAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFG 585

Qy        533 ASASVGYQW 541
              |||||||||
Db        586 ASASVGYQW 594
```

RESULT 13
Q9JPS3
ID Q9JPS3 PRELIMINARY; PRT; 590 AA.
AC Q9JPS3;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Outer membrane protein GNA992.
GN GNA992.
OS Neisseria meningitidis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=487;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NGE28;
RX MEDLINE=20175756; PubMed=10710308;
RA Pizza M., Scarlato V., Masignani V., Giuliani M.M., Arico' B.,
RA Comanducci M., Jennings G.T., Baldi L., Bartolini E., Capecchi B.,
RA Galeotti C.L., Luzzi E., Manetti R., Marchetti E., Mora M., Nuti S.,
RA Ratti G., Santini L., Savino S., Scarselli M., Storni E., Zuo P.,
RA Broeker M., Hundt E., Knapp B., Blair E., Mason T., Tettelin H.,
RA Hood D.W., Jeffries A.C., Saunders N.J., Granoff D.M., Venter C.,
RA Moxon E.R., Grandi G., Rappuoli R.;
RT "Identification of Vaccine Candidates Against Serogroup B
RT Meningococcus by Whole-Genome Sequencing.";
RL Science 287:1816-1820(2000).
DR EMBL; AF226378; AAF42527.1; -.
DR InterPro; IPR005594; YadA.
DR Pfam; PF03895; YadA; 1.
SQ SEQUENCE 590 AA; 61661 MW; 8AA476AC300D80C8 CRC64;

Query Match 87.8%; Score 2401; DB 2; Length 590;
Best Local Similarity 88.5%; Pred. No. 3.1e-92;
Matches 483; Conservative 21; Mismatches 30; Indels 12; Gaps 4;

```
Qy          1 TDEDEEEELESVQRSV-VGSIQASMEGSVELETISLSMTNDSKEFVDPYIV----VTLKA 55
              |||::||:|: |||:| |  : :  ||: | | :    :: :  | :  ::    :||||
Db         52 TDEEDEEDLDPVQRTVAVLIVNSDKEGTGEKEKVE-ENSDWAVYFNEKGVLTAGTITLKA 110

Qy         56 GDNLKIKQNTNENTNASSFTYSLKKDLTGLINVETEKLSFGANGKKVNIISDTKGLNFAK 115
              |||||||      | ::|||||||| | :| |||||| ||| |||| ||||||||||
Db        111 GDNLKIKQ------NGTNFTYSLKKDLTDLTSVGTEKLSFSANGNKVNITSDTKGLNFAK 164

Qy        116 ETAGTNGDTTVHLNGIGSTLTDMLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIK 175
              |||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db        165 ETAGTNGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIK 224

Qy        176 GVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDG 235
              ||||||||||||||||||||||||||||||||||||||||||:||||||||||||||||
Db        225 GVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKRTEVKIGAKTSVIKEKDG 284

Qy        236 KLVTGKGKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGT 295
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        285 KLVTGKGKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGT 344

Qy        296 KVTFASGNGTTATVSKDDQGNITVKYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVIS 355
              |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        345 NVTFASGKGTTATVSKDDQGNITVKYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVIS 404

Qy        356 GNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDE 415
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        405 GNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDE 464

Qy        416 GALNVGSKDANKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVNGNARAGIAQAI 475
              |||||||||||||||||||||||||||||||||||||||||||||||||||:||||||||
Db        465 GALNVGSKDANKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAI 524

Qy        476 ATAGLVQAYLPGKSMMAIGGGTYLGEAGYAIGYSSISAGGNWIIKGTASGNSRGHFGASA 535
              |||||||||||||||||||||| |||||||||||| |||||||||||||||||||||||
Db        525 ATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASA 584

Qy        536 SVGYQW 541
              ||||||
Db        585 SVGYQW 590
```

RESULT 14
Q9JPS6
ID Q9JPS6 PRELIMINARY; PRT; 600 AA.
AC Q9JPS6;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Outer membrane protein GNA992.
GN GNA992.
OS Neisseria meningitidis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=487;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=E26;
RX MEDLINE=20175756; PubMed=10710308;
RA Pizza M., Scarlato V., Masignani V., Giuliani M.M., Arico' B.,
RA Comanducci M., Jennings G.T., Baldi L., Bartolini E., Capecchi B.,
RA Galeotti C.L., Luzzi E., Manetti R., Marchetti E., Mora M., Nuti S.,
RA Ratti G., Santini L., Savino S., Scarselli M., Storni E., Zuo P.,
RA Broeker M., Hundt E., Knapp B., Blair E., Mason T., Tettelin H.,
RA Hood D.W., Jeffries A.C., Saunders N.J., Granoff D.M., Venter C.,
RA Moxon E.R., Grandi G., Rappuoli R.;
RT "Identification of Vaccine Candidates Against Serogroup B
RT Meningococcus by Whole-Genome Sequencing.";
RL Science 287:1816-1820(2000).
DR EMBL; AF226371; AAF42520.1; -.
DR InterPro; IPR005594; YadA.
DR Pfam; PF03895; YadA; 1.
SQ SEQUENCE 600 AA; 62762 MW; 36256963E0598CD1 CRC64;

Query Match 87.6%; Score 2397; DB 2; Length 600;
Best Local Similarity 89.4%; Pred. No. 4.6e-92;
Matches 487; Conservative 12; Mismatches 38; Indels 8; Gaps 4;

```
Qy          2 DEDEEEELESVQRSV-VGSIQASMEGSVELETISLSMTNDSKEFVDPYIV----VTLKAG 56
              | :||| || | |:  | |  :  | : | |       ||   |   ::     :||||
Db         59 DNEEEEYLEPVVRTAPVLSFYSDAEDTGEKEV--TENTNWGIYFDKNGVIKAGTITLKAG 116

Qy         57 DNLKIKQNTNENTNASSFTYSLKKDLTGLINVETEKLSFGANGKKVNIISDTKGLNFAKE 116
              ||||||||:||||||||||||||||:|| | :| ||||||||| |||| |||||||||||
Db        117 DNLKIKQNTDENTNASSFTYSLKKELTDLTSVGTEKLSFGANGNKVNITSDTKGLNFAKE 176

Qy        117 TAGTNGDTTVHLNGIGSTLTDMLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKG 176
              ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        177 TAGTNGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKG 236

Qy        177 VKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGK 236
              |||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||
Db        237 VKPGTTASDNVDFVHTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGK 296

Qy        237 LVTGKGKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTK 296
              ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||| 
Db        297 LVTGKGKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQKGQADKFETVTSGTN 356

Qy        297 VTFASGNGTTATVSKDDQGNITVKYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISG 356
              ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        357 VTFASGKGTTATVSKDDQGNITVKYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISG 416

Qy        357 NVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEG 416
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||  :
Db        417 NVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGD- 475

Qy        417 ALNVGSKDANKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVNGNARAGIAQAIA 476
              ||||||||||||||||||||||||||||||||||||||||||||||:|||||||||||||
Db        476 ALNVGSKDANKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIA 535

Qy        477 TAGLVQAYLPGKSMMAIGGGTYLGEAGYAIGYSSISAGGNWIIKGTASGNSRGHFGASAS 536
              |||||||||||||||||||| ||||||||||||| |||||||||||||||||||||||
Db        536 TAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASAS 595

Qy        537 VGYQW 541
              |||||
Db        596 VGYQW 600
```

```
RESULT 15
Q9JPS4
ID   Q9JPS4      PRELIMINARY;      PRT;   526 AA.
AC   Q9JPS4;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Outer membrane protein GNA992.
GN   GNA992.
OS   Neisseria meningitidis.
OC   Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC   Neisseriaceae; Neisseria.
OX   NCBI_TaxID=487;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NG6/88;
RX   MEDLINE=20175756; PubMed=10710308;
RA   Pizza M., Scarlato V., Masignani V., Giuliani M.M., Arico' B.,
RA   Comanducci M., Jennings G.T., Baldi L., Bartolini E., Capecchi B.,
RA   Galeotti C.L., Luzzi E., Manetti R., Marchetti E., Mora M., Nuti S.,
RA   Ratti G., Santini L., Savino S., Scarselli M., Storni E., Zuo P.,
RA   Broeker M., Hundt E., Knapp B., Blair E., Mason T., Tettelin H.,
RA   Hood D.W., Jeffries A.C., Saunders N.J., Granoff D.M., Venter C.,
RA   Moxon E.R., Grandi G., Rappuoli R.;
RT   "Identification of Vaccine Candidates Against Serogroup B
RT   Meningococcus by Whole-Genome Sequencing.";
RL   Science 287:1816-1820(2000).
DR   EMBL; AF226377; AAF42526.1; -.
DR   InterPro; IPR005594; YadA.
DR   Pfam; PF03895; YadA; 1.
SQ   SEQUENCE   526 AA;  54732 MW;  F543C4234AF807CA CRC64;
```

```
  Query Match             86.4%;  Score 2363.5;  DB 2;  Length 526;
  Best Local Similarity   95.2%;  Pred. No. 9.6e-91;
  Matches  471;  Conservative    5;  Mismatches   14;  Indels    5;  Gaps    2;

Qy         51 VTLKAGDNLKIKQ----NTNENTNASSFTYSLKKDLTGLINVETEKLSFGANGKKVNIIS 106
              |||||||||||||    |||||| |||||||||||| | :||||||||||| |||| |
Db         33 VTLKAGDNLKIKQNTNKNTNENTNDSSFTYSLKKDLTDLTSVETEKLSFGANGNKVNITS 92

Qy        107 DTKGLNFAKETAGTNGDTTVHLNGIGSTLTDMLLNTGATTNVTNDNVTDDEKKRAASVKD 166
              ||||||||||||||||||||||||||||||| ||||||||||||||||||:|||||||||
Db         93 DTKGLNFAKETAGTNGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDKKKRAASVKD 152

Qy        167 VLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAK 226
              |||||||||||||||||||||||| ||||||||||||||||||||||||||:|||||||||
Db        153 VLNAGWNIKGVKPGTTASDNVDFVHTYDTVEFLSADTKTTTVNVESKDNGKRTEVKIGAK 212

Qy        227 TSVIKEKDGKLVTGKGKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQAD 286
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        213 TSVIKEKDGKLVTGKGKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQAD 272

Qy        287 KFETVTSGTKVTFASGNGTTATVSKDDQGNITVKYDVNVGDALNVNQLQNSGWNLDSKAV 346
              |||||||| ||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        273 KFETVTSGTNVTFASGKGTTATVSKDDQGNITVKYDVNVGDALNVNQLQNSGWNLDSKAV 332

Qy        347 AGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGAD 406
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        333 AGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGAD 392

Qy        407 APTLSVDDEGALNVGSKDANKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVNGN 466
              |||||| : |||||| |||||||||||||||||||||||||||||||||||||||||:||
Db        393 APTLSVDGD-ALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGN 451

Qy        467 ARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYLGEAGYAIGYSSISAGGNWIIKGTASGN 526
              ||||||||||||||||||||||||||||||| ||||||||||||| ||||||||||||||
Db        452 ARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGN 511

Qy        527 SRGHFGASASVGYQW 541
              |||||||||||||||
Db        512 SRGHFGASASVGYQW 526


Search completed: October  6, 2003, 09:30:48
Job time : 40.7345 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: October 6, 2003, 09:05:45 ; Search time 39.6746 Seconds
(without alignments)
1848.329 Million cell updates/sec

Title: US-09-771-382-36
Perfect score: 2350
Sequence: 1 TDETGLINVETEKLSFGANG.....TASGNSRGHFGASASVGYQW 462

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03:*
1: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2350	100.0	513	22	AAU06183	N. meningitidis H4
2	2337	99.4	592	20	AAY23744	A surface protein
3	2337	99.4	592	22	AAU06172	N. meningitidis H4
4	2261	96.2	598	20	AAY23738	A surface protein
5	2261	96.2	598	20	AAY23742	A surface protein
6	2261	96.2	598	22	AAU06177	N. meningitidis H1
7	2261	96.2	598	22	AAU06178	N. meningitidis BZ
8	2260	96.2	599	20	AAY23743	A surface protein
9	2260	96.2	599	22	AAU06176	N. meningitidis H3
10	2247	95.6	594	20	AAY23739	A surface protein
11	2247	95.6	594	22	AAU06179	N. meningitidis BZ
12	2213.5	94.2	512	22	AAU06182	N. meningitidis PM
13	2213	94.2	594	20	AAY23740	A surface protein
14	2213	94.2	594	21	AAY57044	BASB029 amino acid
15	2213	94.2	594	22	AAU06174	N. meningitidis EG
16	2210.5	94.1	591	20	AAY27202	Amino acid sequenc
17	2210.5	94.1	591	20	AAY23741	A surface protein
18	2210.5	94.1	591	20	AAY23746	A surface protein
19	2210.5	94.1	591	22	AAU06171	N. meningitidis PM
20	2210.5	94.1	591	22	AAU06175	N. meningitidis EG
21	2207.5	93.9	591	21	AAY57045	BASB029 amino acid
22	2206.5	93.9	592	20	AAY23737	A surface protein
23	2136	90.9	592	22	AAU06180	N. meningitidis Z2
24	2116	90.0	589	20	AAY23745	A surface protein
25	2116	90.0	589	22	AAU06173	N. meningitidis P2
26	2113	89.9	592	20	AAY27203	Amino acid sequenc
27	2032	86.5	502	22	AAU06186	N. meningitidis PM
28	1953	83.1	604	22	AAU06181	N. meningitidis su
29	1861.5	79.2	433	22	AAU06185	N. meningitidis PM
30	1740.5	74.1	407	22	AAU06184	N. meningitidis PM
31	876	37.3	2353	17	AAR99393	Haemophilus adhesi
32	876	37.3	2411	21	AAB23860	Haemophilus influe
33	865	36.8	1094	21	AAB23858	Haemophilus influe
34	849.5	36.1	1098	17	AAR99392	Haemophilus adhesi
35	727	30.9	245	20	AAY27201	Amino acid sequenc
36	613	26.1	679	17	AAR99394	Haemophilus adhesi
37	613	26.1	679	21	AAB23855	Haemophilus influe
38	580	24.7	116	21	AAB37832	Neisserial conserv
39	571.5	24.3	1002	21	AAB23854	Haemophilus influe
40	551.5	23.5	1004	21	AAB23857	Haemophilus influe
41	511	21.7	1104	21	AAB23856	Haemophilus influe
42	511	21.7	1104	21	AAB23859	Haemophilus influe
43	390.5	16.6	298	24	AAE30477	Haemophilus influe
44	371.5	15.8	2139	24	ABP71294	M. catarrhalis sur
45	371	15.8	2314	22	AAB69136	M. catarrhalis les

ALIGNMENTS

RESULT 1
AAU06183
ID AAU06183 standard; Protein; 513 AA.
XX
AC AAU06183;
XX
DT 24-OCT-2001 (first entry)
XX
DE N. meningitidis H41 NhhA deletion mutant.
XX
KW Surface antigen NhhA; meningococcal disease; meningitis vaccine;
KW mutant; mutein.
XX
OS Neisseria meningitidis strain H41.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Peptide 1..51
FT /label= Signal_peptide
FT Protein 52..513
FT /label= Mature_NhhA_deletion_mutant
FT /note= "Predicted mature protein, specifically
FT claimed in claim 12"
XX
PN WO200155182-A1.
XX
PD 02-AUG-2001.
XX
PF 25-JAN-2001; 2001WO-AU00069.
XX
PR 25-JAN-2000; 2000US-0177917.
XX

```
PA   (UYQU ) UNIV QUEENSLAND.
XX
PI   Peak IRA, Jennings MP;
XX
DR   WPI; 2001-488774/53.
DR   N-PSDB; AAS09173.
XX
PT   New NhhA surface antigen polypeptides and polynucleotides from
PT   Neisseria meningitidis, useful in producing vaccines for treating or
PT   preventing broad spectrum of Neisseria meningitidis -
XX
PS   Claim 12; Fig 6; 91pp; English.
XX
CC   The present invention relates to the isolation of novel Neisseria
CC   meningitidis mutant polypeptides of the surface antigen NhhA
CC   (AAU06182-AAU06186). The modified or mutant NhhA polypeptides are
CC   characterised by deletions of non-conserved amino acids, particularly
CC   the deletion of variable regions. The deletion mutants are useful in
CC   diagnostics, therapeutic and prophylactic vaccines against a broader
CC   spectrum of N. meningitidis, and in designing and/or screening of
CC   medicaments. The mutant proteins when used as a vaccine can effectively
CC   immunise against a broader spectrum of N. meningitidis strains than
CC   would be expected from a corresponding wild-type surface antigen.
CC   The present sequence represents N. meningitidis strain H41 surface
CC   antigen NhhA deletion mutant.
XX
SQ   Sequence   513 AA;

  Query Match            100.0%;  Score 2350;  DB 22;  Length 513;
  Best Local Similarity  100.0%;  Pred. No. 1e-139;
  Matches  462;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TDETGLINVETEKLSFGANGKKVNIISDTKGLNFAKETAGTNGDTTVHLNGIGSTLTDML 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         52 TDETGLINVETEKLSFGANGKKVNIISDTKGLNFAKETAGTNGDTTVHLNGIGSTLTDML 111

Qy         61 LNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        112 LNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFL 171

Qy        121 SADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKGKGENGSSTDEGEGLVTA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        172 SADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKGKGENGSSTDEGEGLVTA 231

Qy        181 KEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTKVTFASGNGTTATVSKDDQGNITV 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        232 KEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTKVTFASGNGTTATVSKDDQGNITV 291

Qy        241 KYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEI 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        292 KYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEI 351

Qy        301 TRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGALNVGSKDANKPVRITNVAPGVK 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        352 TRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGALNVGSKDANKPVRITNVAPGVK 411

Qy        361 EGDVTNVAQLKGVAQNLNNRIDNVNGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYL 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        412 EGDVTNVAQLKGVAQNLNNRIDNVNGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYL 471

Qy        421 GEAGYAIGYSSISAGGNWIIKGTASGNSRGHFGASASVGYQW 462
              |||||||||||||||||||||||||||||||||||||||||||
Db        472 GEAGYAIGYSSISAGGNWIIKGTASGNSRGHFGASASVGYQW 513


RESULT 2
AAY23744
ID   AAY23744 standard; Protein; 592 AA.
XX
AC   AAY23744;
XX
DT   08-SEP-1999  (first entry)
XX
DE   A surface protein of Neisseria meningitidis.
XX
KW   Surface protein; surface glycoprotein; infection; vaccine;
KW   immunoreactive peptide.
XX
OS   Neisseria meningitidis.
XX
PN   WO9931132-A1.
XX
PD   24-JUN-1999.
XX
PF   14-DEC-1998;   98WO-AU01031.
XX
PR   12-DEC-1997;   97GB-0026398.
XX
PA   (ISIS-) ISIS INNOVATION LTD.
PA   (UYQU ) UNIV QUEENSLAND.
XX
PI   Jennings MP,  Moxon ER,  Peak IRA;
XX
DR   WPI; 1999-418754/35.
DR   N-PSDB; AAX85796.
XX
PT   Neisseria meningitidis surface proteins useful for treating N.
PT   meningitidis infections
XX
PS   Claim 1; Page 118-120; 132pp; English.
XX
CC   The present sequence represents a surface protein of Neiserria
CC   meningitidis which is approximately 62 kDa. The N. meningitidis
CC   surface glycoproteins, nucleic acids, the primers and optionally
CC   a thermostable polymerase, or antibodies are useful in a kit for
CC   the detection or diagnosis of N. meningitidis infection in humans.
CC   The N. meningitidis surface glycoproteins can also be used to
CC   prevent or treat N. meningitidis infection in humans, especially
CC   in the form of vaccines. The proteins and antibodies can also
CC   be used to identify immunoreactive peptides.
XX
SQ   Sequence   592 AA;

  Query Match            99.4%;  Score 2337;  DB 20;  Length 592;
  Best Local Similarity  99.8%;  Pred. No. 8.1e-139;
  Matches  460;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 DETGLINVETEKLSFGANGKKVNIISDTKGLNFAKETAGTNGDTTVHLNGIGSTLTDMLL 61
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        132 DLTGLINVETEKLSFGANGKKVNIISDTKGLNFAKETAGTNGDTTVHLNGIGSTLTDMLL 191

Qy         62 NTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLS 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        192 NTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLS 251

Qy        122 ADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKGKGENGSSTDEGEGLVTAK 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        252 ADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKGKGENGSSTDEGEGLVTAK 311

Qy        182 EVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTKVTFASGNGTTATVSKDDQGNITVK 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        312 EVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTKVTFASGNGTTATVSKDDQGNITVK 371

Qy        242 YDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEIT 301
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        372 YDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEIT 431

Qy        302 RNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGALNVGSKDANKPVRITNVAPGVKE 361
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        432 RNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGALNVGSKDANKPVRITNVAPGVKE 491

Qy        362 GDVTNVAQLKGVAQNLNNRIDNVNGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYLG 421
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        492 GDVTNVAQLKGVAQNLNNRIDNVNGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYLG 551

Qy        422 EAGYAIGYSSISAGGNWIIKGTASGNSRGHFGASASVGYQW 462
```

```
                    ||||||||||||||||||||||||||||||||||||||||
Db        552 EAGYAIGYSSISAGGNWIIKGTASGNSRGHFGASASVGYQW 592


RESULT 3
AAU06172
ID   AAU06172 standard; Protein; 592 AA.
XX
AC   AAU06172;
XX
DT   24-OCT-2001  (first entry)
XX
DE   N. meningitidis H41 surface antigen NhhA polypeptide sequence.
XX
KW   Surface antigen NhhA; meningococcal disease; meningitis vaccine.
XX
OS   Neisseria meningitidis strain H41.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..51
FT                   /label= Signal_peptide
FT   Region          1..50
FT                   /label= C1
FT                   /note= "Conserved region 1"
FT   Region          51..102
FT                   /label= V1
FT                   /note= "Variable region 1"
FT   Protein         52..592
FT                   /label= Mature_NhhA
FT                   /note= "Predicted mature protein, specifically
FT                        claimed in claim 12"
FT   Region          103..114
FT                   /label= C2
FT                   /note= "Conserved region 2"
FT   Region          115..124
FT                   /label= V2
FT                   /note= "Variable region 2"
FT   Region          125..188
FT                   /label= C3
FT                   /note= "Conserved region 3"
FT   Region          189..210
FT                   /label= V3
FT                   /note= "Variable region 3"
FT   Region          211..229
FT                   /label= C4
FT                   /note= "Conserved region 4"
FT   Region          230..236
FT                   /label= V4
FT                   /note= "Variable region 4"
FT   Region          237..592
FT                   /label= C5
FT                   /note= "Conserved region 5"
XX
PN   WO200155182-A1.
XX
PD   02-AUG-2001.
XX
PF   25-JAN-2001; 2001WO-AU00069.
XX
PR   25-JAN-2000; 2000US-0177917.
XX
PA   (UYQU ) UNIV QUEENSLAND.
XX
PI   Peak IRA,  Jennings MP;
XX
DR   WPI; 2001-488774/53.
DR   N-PSDB; AAS09162.
XX
PT   New NhhA surface antigen polypeptides and polynucleotides from
PT   Neisseria meningitidis, useful in producing vaccines for treating or
PT   preventing broad spectrum of Neisseria meningitidis -
XX
PS   Claim 9; Fig 1; 91pp; English.
XX
CC   The present invention relates to the isolation of novel Neisseria
CC   meningitidis mutant polypeptides of the surface antigen NhhA
CC   (AAU06182-AAU06186). The modified or mutant NhhA polypeptides are
CC   characterised by deletions of non-conserved amino acids, particularly
CC   the deletion of variable regions. The deletion mutants are useful in
CC   diagnostics, therapeutic and prophylactic vaccines against a broader
CC   spectrum of N. meningitidis, and in designing and/or screening of
CC   medicaments. The mutant proteins when used as a vaccine can effectively
CC   immunise against a broader spectrum of N. meningitidis strains than
CC   would be expected from a corresponding wild-type surface antigen.
CC   The present sequence representing the wild type surface antigen NhhA
CC   from N. meningitidis strain H41 is 1 of 10 NhhA polypeptide sequences
CC   (AAU06171-AAU06180) from 10 different N. meningitidis strains given in
CC   the present invention.
XX
SQ   Sequence   592 AA;

  Query Match             99.4%;  Score 2337;  DB 22;  Length 592;
  Best Local Similarity   99.8%;  Pred. No. 8.1e-139;
  Matches  460;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 DETGLINVETEKLSFGANGKKVNIISDTKGLNFAKETAGTNGDTTVHLNGIGSTLTDMLL 61
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        132 DLTGLINVETEKLSFGANGKKVNIISDTKGLNFAKETAGTNGDTTVHLNGIGSTLTDMLL 191

Qy         62 NTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLS 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        192 NTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLS 251

Qy        122 ADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKGKGENGSSTDEGEGLVTAK 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        252 ADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKGKGENGSSTDEGEGLVTAK 311

Qy        182 EVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTKVTFASGNGTTATVSKDDQGNITVK 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        312 EVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTKVTFASGNGTTATVSKDDQGNITVK 371

Qy        242 YDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEIT 301
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        372 YDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEIT 431

Qy        302 RNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGALNVGSKDANKPVRITNVAPGVKE 361
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        432 RNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGALNVGSKDANKPVRITNVAPGVKE 491

Qy        362 GDVTNVAQLKGVAQNLNNRIDNVNGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYLG 421
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        492 GDVTNVAQLKGVAQNLNNRIDNVNGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYLG 551

Qy        422 EAGYAIGYSSISAGGNWIIKGTASGNSRGHFGASASVGYQW 462
              |||||||||||||||||||||||||||||||||||||||||
Db        552 EAGYAIGYSSISAGGNWIIKGTASGNSRGHFGASASVGYQW 592


RESULT 4
AAY23738
ID   AAY23738 standard; Protein; 598 AA.
XX
AC   AAY23738;
XX
DT   08-SEP-1999  (first entry)
XX
DE   A surface protein of Neisseria meningitidis.
XX
KW   Surface protein; surface glycoprotein; infection; vaccine;
KW   immunoreactive peptide.
XX
OS   Neisseria meningitidis.
XX
PN   WO9931132-A1.
XX
PD   24-JUN-1999.
```

```
XX
PF   14-DEC-1998;   98WO-AU01031.
XX
PR   12-DEC-1997;   97GB-0026398.
XX
PA   (ISIS-) ISIS INNOVATION LTD.
PA   (UYQU ) UNIV QUEENSLAND.
XX
PI   Jennings MP,  Moxon ER,  Peak IRA;
XX
DR   WPI; 1999-418754/35.
DR   N-PSDB; AAX85790.
XX
PT   Neisseria meningitidis surface proteins useful for treating N.
PT   meningitidis infections
XX
PS   Claim 1; Page 91-93; 132pp; English.
XX
CC   The present sequence represents a surface protein of Neiserria
CC   meningitidis which is approximately 62 kDa. The N. meningitidis
CC   surface glycoproteins, nucleic acids, the primers and optionally
CC   a thermostable polymerase, or antibodies are useful in a kit for
CC   the detection or diagnosis of N. meningitidis infection in humans.
CC   The N. meningitidis surface glycoproteins can also be used to
CC   prevent or treat N. meningitidis infection in humans, especially
CC   in the form of vaccines. The proteins and antibodies can also
CC   be used to identify immunoreactive peptides.
XX
SQ   Sequence   598 AA;

  Query Match             96.2%;  Score 2261;  DB 20;  Length 598;
  Best Local Similarity   96.5%;  Pred. No. 4.9e-134;
  Matches  445;  Conservative    4;  Mismatches   12;  Indels    0;  Gaps    0;

Qy          2 DETGLINVETEKLSFGANGKKVNIISDTKGLNFAKETAGTNGDTTVHLNGIGSTLTDMLL 61
              | | | :||||||||||||| |||| |||||||||||||||||| |||||||||||| ||
Db        138 DLTDLTSVETEKLSFGANGNKVNITSDTKGLNFAKETAGTNGDPTVHLNGIGSTLTDTLL 197

Qy         62 NTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLS 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        198 NTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLS 257

Qy        122 ADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKGKGENGSSTDEGEGLVTAK 181
              ||||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||
Db        258 ADTKTTTVNVESKDNGKRTEVKIGAKTSVIKEKDGKLVTGKGKGENGSSTDEGEGLVTAK 317

Qy        182 EVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTKVTFASGNGTTATVSKDDQGNITVK 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        318 EVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTKVTFASGNGTTATVSKDDQGNITVK 377

Qy        242 YDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEIT 301
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        378 YDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEIT 437

Qy        302 RNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGALNVGSKDANKPVRITNVAPGVKE 361
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        438 RNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGALNVGSKDANKPVRITNVAPGVKE 497

Qy        362 GDVTNVAQLKGVAQNLNNRIDNVNGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYLG 421
              |||||||||||||||||||||:|||||||||||||||| |||||||||||||||||| |
Db        498 GDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLAQAYLPGKSMMAIGGGTYRG 557

Qy        422 EAGYAIGYSSISAGGNWIIKGTASGNSRGHFGASASVGYQW 462
              |||||||||||  |||:|||||||||||||| ||||||||
Db        558 EAGYAIGYSSISDTGNWVIKGTASGNSRGHFGTSASVGYQW 598


RESULT 5
AAY23742
ID   AAY23742 standard; Protein; 598 AA.
XX
AC   AAY23742;
XX
DT   08-SEP-1999  (first entry)
XX
DE   A surface protein of Neisseria meningitidis.
XX
KW   Surface protein; surface glycoprotein; infection; vaccine;
KW   immunoreactive peptide.
XX
OS   Neisseria meningitidis.
XX
PN   WO9931132-A1.
XX
PD   24-JUN-1999.
XX
PF   14-DEC-1998;   98WO-AU01031.
XX
PR   12-DEC-1997;   97GB-0026398.
XX
PA   (ISIS-) ISIS INNOVATION LTD.
PA   (UYQU ) UNIV QUEENSLAND.
XX
PI   Jennings MP,  Moxon ER,  Peak IRA;
XX
DR   WPI; 1999-418754/35.
DR   N-PSDB; AAX85794.
XX
PT   Neisseria meningitidis surface proteins useful for treating N.
PT   meningitidis infections
XX
PS   Claim 1; Page 108-110; 132pp; English.
XX
CC   The present sequence represents a surface protein of Neiserria
CC   meningitidis which is approximately 62 kDa. The N. meningitidis
CC   surface glycoproteins, nucleic acids, the primers and optionally
CC   a thermostable polymerase, or antibodies are useful in a kit for
CC   the detection or diagnosis of N. meningitidis infection in humans.
CC   The N. meningitidis surface glycoproteins can also be used to
CC   prevent or treat N. meningitidis infection in humans, especially
CC   in the form of vaccines. The proteins and antibodies can also
CC   be used to identify immunoreactive peptides.
XX
SQ   Sequence   598 AA;

  Query Match             96.2%;  Score 2261;  DB 20;  Length 598;
  Best Local Similarity   96.7%;  Pred. No. 4.9e-134;
  Matches  446;  Conservative    3;  Mismatches   12;  Indels    0;  Gaps    0;

Qy          2 DETGLINVETEKLSFGANGKKVNIISDTKGLNFAKETAGTNGDTTVHLNGIGSTLTDMLL 61
              | | | :||||||||||||| |||| |||||||||||||||||| |||||||||||| ||
Db        138 DLTDLTSVETEKLSFGANGNKVNITSDTKGLNFAKETAGTNGDPTVHLNGIGSTLTDTLL 197

Qy         62 NTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLS 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        198 NTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLS 257

Qy        122 ADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKGKDENGSSTDEGEGLVTAK 181
              |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        258 ADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKGKDENGSSTDEGEGLVTAK 317

Qy        182 EVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTKVTFASGNGTTATVSKDDQGNITVK 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        318 EVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTKVTFASGNGTTATVSKDDQGNITVK 377

Qy        242 YDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEIT 301
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        378 YDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEIT 437

Qy        302 RNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGALNVGSKDANKPVRITNVAPGVKE 361
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        438 RNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGALNVGSKDANKPVRITNVAPGVKE 497

Qy        362 GDVTNVAQLKGVAQNLNNRIDNVNGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYLG 421
              |||||||||||||||||||||:|||||||||||||||| |||||||||||||||||| |
Db        498 GDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLAQAYLPGKSMMAIGGGTYRG 557
```

```
Qy        422 EAGYAIGYSSISAGGNWIIKGTASGNSRGHFGASASVGYQW 462
              ||||||||||||  |||:||||||||||||||||||||||
Db        558 EAGYAIGYSSISDTGNWVIKGTASGNSRGHFGASASVGYQW 598


RESULT 6
AAU06177
ID   AAU06177 standard; Protein; 598 AA.
XX
AC   AAU06177;
XX
DT   24-OCT-2001  (first entry)
XX
DE   N. meningitidis H15 surface antigen NhhA polypeptide sequence.
XX
KW   Surface antigen NhhA; meningococcal disease; meningitis vaccine.
XX
OS   Neisseria meningitidis strain H15.
XX
FH   Key             Location/Qualifiers
FT   Region          1..50
FT                   /label= C1
FT                   /note= "Conserved region 1"
FT   Region          51..104
FT                   /label= V1
FT                   /note= "Variable region 1"
FT   Region          105..116
FT                   /label= C2
FT                   /note= "Conserved region 2"
FT   Region          117..130
FT                   /label= V2
FT                   /note= "Variable region 2"
FT   Region          131..194
FT                   /label= C3
FT                   /note= "Conserved region 3"
FT   Region          195..216
FT                   /label= V3
FT                   /note= "Variable region 3"
FT   Region          217..235
FT                   /label= C4
FT                   /note= "Conserved region 4"
FT   Region          236..242
FT                   /label= V4
FT                   /note= "Variable region 4"
FT   Region          243..598
FT                   /label= C5
FT                   /note= "Conserved region 5"
XX
PN   WO200155182-A1.
XX
PD   02-AUG-2001.
XX
PF   25-JAN-2001; 2001WO-AU00069.
XX
PR   25-JAN-2000; 2000US-0177917.
XX
PA   (UYQU ) UNIV QUEENSLAND.
XX
PI   Peak IRA,  Jennings MP;
XX
DR   WPI; 2001-488774/53.
DR   N-PSDB; AAS09167.
XX
PT   New NhhA surface antigen polypeptides and polynucleotides from
PT   Neisseria meningitidis, useful in producing vaccines for treating or
PT   preventing broad spectrum of Neisseria meningitidis -
XX
PS   Claim 9; Fig 1; 91pp; English.
XX
CC   The present invention relates to the isolation of novel Neisseria
CC   meningitidis mutant polypeptides of the surface antigen NhhA
CC   (AAU06182-AAU06186). The modified or mutant NhhA polypeptides are
CC   characterised by deletions of non-conserved amino acids, particularly
CC   the deletion of variable regions. The deletion mutants are useful in
CC   diagnostics, therapeutic and prophylactic vaccines against a broader
CC   spectrum of N. meningitidis, and in designing and/or screening of
CC   medicaments. The mutant proteins when used as a vaccine can effectively
CC   immunise against a broader spectrum of N. meningitidis strains than
CC   would be expected from a corresponding wild-type surface antigen.
CC   The present sequence representing the wild type surface antigen NhhA
CC   from N. meningitidis strain H15 is 1 of 10 NhhA polypeptide sequences
CC   (AAU06171-AAU06180) from 10 different N. meningitidis strains given in
CC   the present invention.
XX
SQ   Sequence   598 AA;

  Query Match             96.2%;  Score 2261;  DB 22;  Length 598;
  Best Local Similarity   96.7%;  Pred. No. 4.9e-134;
  Matches  446;  Conservative    3;  Mismatches   12;  Indels    0;  Gaps    0;

Qy          2 DETGLINVETEKLSFGANGKKVNIISDTKGLNFAKETAGTNGDTTVHLNGIGSTLTDMLL 61
              | | | :||||||||||||| |||| |||||||||||||||||| |||||||||||| ||
Db        138 DLTDLTSVETEKLSFGANGNKVNITSDTKGLNFAKETAGTNGDPTVHLNGIGSTLTDTLL 197

Qy         62 NTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLS 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        198 NTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLS 257

Qy        122 ADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKGKGENGSSTDEGEGLVTAK 181
              |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        258 ADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKGKDENGSSTDEGEGLVTAK 317

Qy        182 EVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTKVTFASGNGTTATVSKDDQGNITVK 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        318 EVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTKVTFASGNGTTATVSKDDQGNITVK 377

Qy        242 YDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEIT 301
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        378 YDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEIT 437

Qy        302 RNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGALNVGSKDANKPVRITNVAPGVKE 361
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        438 RNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGALNVGSKDANKPVRITNVAPGVKE 497

Qy        362 GDVTNVAQLKGVAQNLNNRIDNVGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYLG 421
              |||||||||||||||||||||:||||||||||||||| |||||||||||||||| |
Db        498 GDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLAQAYLPGKSMMAIGGGTYRG 557

Qy        422 EAGYAIGYSSISAGGNWIIKGTASGNSRGHFGASASVGYQW 462
              ||||||||||||  |||:||||||||||||||||||||||
Db        558 EAGYAIGYSSISDTGNWVIKGTASGNSRGHFGASASVGYQW 598


RESULT 7
AAU06178
ID   AAU06178 standard; Protein; 598 AA.
XX
AC   AAU06178;
XX
DT   24-OCT-2001  (first entry)
XX
DE   N. meningitidis BZ10 surface antigen NhhA polypeptide sequence.
XX
KW   Surface antigen NhhA; meningococcal disease; meningitis vaccine.
XX
OS   Neisseria meningitidis strain BZ10.
XX
FH   Key             Location/Qualifiers
FT   Region          1..50
FT                   /label= C1
FT                   /note= "Conserved region 1"
FT   Region          51..104
FT                   /label= V1
FT                   /note= "Variable region 1"
FT   Region          105..116
```

```
FT                   /label= C2
FT                   /note= "Conserved region 2"
FT   Region          117..130
FT                   /label= V2
FT                   /note= "Variable region 2"
FT   Region          131..194
FT                   /label= C3
FT                   /note= "Conserved region 3"
FT   Region          195..216
FT                   /label= V3
FT                   /note= "Variable region 3"
FT   Region          217..235
FT                   /label= C4
FT                   /note= "Conserved region 4"
FT   Region          236..242
FT                   /label= V4
FT                   /note= "Variable region 4"
FT   Region          243..598
FT                   /label= C5
FT                   /note= "Conserved region 5"
XX
PN   WO200155182-A1.
XX
PD   02-AUG-2001.
XX
PF   25-JAN-2001; 2001WO-AU00069.
XX
PR   25-JAN-2000; 2000US-0177917.
XX
PA   (UYQU ) UNIV QUEENSLAND.
XX
PI   Peak IRA,  Jennings MP;
XX
DR   WPI; 2001-488774/53.
DR   N-PSDB; AAS09168.
XX
PT   New NhhA surface antigen polypeptides and polynucleotides from
PT   Neisseria meningitidis, useful in producing vaccines for treating or
PT   preventing broad spectrum of Neisseria meningitidis -
XX
PS   Claim 9; Fig 1; 91pp; English.
XX
CC   The present invention relates to the isolation of novel Neisseria
CC   meningitidis mutant polypeptides of the surface antigen NhhA
CC   (AAU06182-AAU06186). The modified or mutant NhhA polypeptides are
CC   characterised by deletions of non-conserved amino acids, particularly
CC   the deletion of variable regions. The deletion mutants are useful in
CC   diagnostics, therapeutic and prophylactic vaccines against a broader
CC   spectrum of N. meningitidis, and in designing and/or screening of
CC   medicaments. The mutant proteins when used as a vaccine can effectively
CC   immunise against a broader spectrum of N. meningitidis strains than
CC   would be expected from a corresponding wild-type surface antigen.
CC   The present sequence representing the wild type surface antigen NhhA
CC   from N. meningitidis strain BZ10 is 1 of 10 NhhA polypeptide sequences
CC   (AAU06171-AAU06180) from 10 different N. meningitidis strains given in
CC   the present invention.
XX
SQ   Sequence   598 AA;

  Query Match            96.2%;  Score 2261;  DB 22;  Length 598;
  Best Local Similarity  96.5%;  Pred. No. 4.9e-134;
  Matches  445;  Conservative    4;  Mismatches   12;  Indels    0;  Gaps    0;

Qy          2 DETGLINVETEKLSFGANGKKVNIISDTKGLNFAKETAGTNGDTTVHLNGIGSTLTDMLL 61
              | | | :||||||||||||| |||| |||||||||||||||||| |||||||||||| ||
Db        138 DLTDLTSVETEKLSFGANGNKVNITSDTKGLNFAKETAGTNGDPTVHLNGIGSTLTDTLL 197

Qy         62 NTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLS 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        198 NTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLS 257

Qy        122 ADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKGKGENGSSTDEGEGLVTAK 181
              ||||||||||||||||:||||||||||||||||||||||||||||||||||||||||||||

Db        258 ADTKTTTVNVESKDNGKRTEVKIGAKTSVIKEKDGKLVTGKGKGENGSSTDEGEGLVTAK 317

Qy        182 EVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTKVTFASGNGTTATVSKDDQGNITVK 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        318 EVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTKVTFASGNGTTATVSKDDQGNITVK 377

Qy        242 YDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEIT 301
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        378 YDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEIT 437

Qy        302 RNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGALNVGSKDANKPVRITNVAPGVKE 361
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        438 RNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGALNVGSKDANKPVRITNVAPGVKE 497

Qy        362 GDVTNVAQLKGVAQNLNNRIDNVNGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYLG 421
              ||||||||||||||||||||:|||||||||||||||| |||||||||||||||| |
Db        498 GDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLAQAYLPGKSMMAIGGGTYRG 557

Qy        422 EAGYAIGYSSISAGGNWIIKGTASGNSRGHFGASASVGYQW 462
              |||||||||||  |||:||||||||||||| ||||||||
Db        558 EAGYAIGYSSISDTGNWVIKGTASGNSRGHFGTSASVGYQW 598


RESULT 8
AAY23743
ID   AAY23743 standard; Protein; 599 AA.
XX
AC   AAY23743;
XX
DT   08-SEP-1999  (first entry)
XX
DE   A surface protein of Neisseria meningitidis.
XX
KW   Surface protein; surface glycoprotein; infection; vaccine;
KW   immunoreactive peptide.
XX
OS   Neisseria meningitidis.
XX
PN   WO9931132-A1.
XX
PD   24-JUN-1999.
XX
PF   14-DEC-1998;   98WO-AU01031.
XX
PR   12-DEC-1997;   97GB-0026398.
XX
PA   (ISIS-) ISIS INNOVATION LTD.
PA   (UYQU ) UNIV QUEENSLAND.
XX
PI   Jennings MP,  Moxon ER,  Peak IRA;
XX
DR   WPI; 1999-418754/35.
DR   N-PSDB; AAX85795.
XX
PT   Neisseria meningitidis surface proteins useful for treating N.
PT   meningitidis infections
XX
PS   Claim 1; Page 114-115; 132pp; English.
XX
CC   The present sequence represents a surface protein of Neiserria
CC   meningitidis which is approximately 62 kDa. The N. meningitidis
CC   surface glycoproteins, nucleic acids, the primers and optionally
CC   a thermostable polymerase, or antibodies are useful in a kit for
CC   the detection or diagnosis of N. meningitidis infection in humans.
CC   The N. meningitidis surface glycoproteins can also be used to
CC   prevent or treat N. meningitidis infection in humans, especially
CC   in the form of vaccines. The proteins and antibodies can also
CC   be used to identify immunoreactive peptides.
XX
SQ   Sequence   599 AA;

  Query Match            96.2%;  Score 2260;  DB 20;  Length 599;
  Best Local Similarity  96.5%;  Pred. No. 5.6e-134;
```

```
  Matches  445;  Conservative    5;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          2 DETGLINVETEKLSFGANGKKVNIISDTKGLNFAKETAGTNGDTTVHLNGIGSTLTDMLL 61
              | | | :||||||||||| |||| ||||||||||||||||||||||||||||||||| ||
Db        139 DLTDLTSVETEKLSFGANGNKVNITSDTKGLNFAKETAGTNGDTTVHLNGIGSTLTDTLL 198

Qy         62 NTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLS 121
              ||||||||||||||||:|||||||||||||||||||||||||||||||||| |||||||||
Db        199 NTGATTNVTNDNVTDDKKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVHTYDTVEFLS 258

Qy        122 ADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKGKGENGSSTDEGEGLVTAK 181
              ||||||||||||||||:||||||||||||||||||||||||||||||||||||||||||||
Db        259 ADTKTTTVNVESKDNGKRTEVKIGAKTSVIKEKDGKLVTGKGKGENGSSTDEGEGLVTAK 318

Qy        182 EVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTKVTFASGNGTTATVSKDDQGNITVK 241
              |||||||||||||||||||||||||||||||||| |||||| ||||||||||||||||||
Db        319 EVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVK 378

Qy        242 YDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEIT 301
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        379 YDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEIT 438

Qy        302 RNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGALNVGSKDANKPVRITNVAPGVKE 361
              |||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||
Db        439 RNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDKGALNVGSKDANKPVRITNVAPGVKE 498

Qy        362 GDVTNVAQLKGVAQNLNNRIDNVNGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYLG 421
              |||||||||||||||||||||:|||||||||||||||||||||||||||||||||||| |
Db        499 GDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRG 558

Qy        422 EAGYAIGYSSISAGGNWIIKGTASGNSRGHFGASASVGYQW 462
              |||||||||||| ||||||||||||||||||||||||||||
Db        559 EAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 599


RESULT 9
AAU06176
ID   AAU06176 standard; Protein; 599 AA.
XX
AC   AAU06176;
XX
DT   24-OCT-2001  (first entry)
XX
DE   N. meningitidis H38 surface antigen NhhA polypeptide sequence.
XX
KW   Surface antigen NhhA; meningococcal disease; meningitis vaccine.
XX
OS   Neisseria meningitidis strain H38.
XX
FH   Key             Location/Qualifiers
FT   Region          1..50
FT                   /label= C1
FT                   /note= "Conserved region 1"
FT   Region          51..105
FT                   /label= V1
FT                   /note= "Variable region 1"
FT   Region          106..117
FT                   /label= C2
FT                   /note= "Conserved region 2"
FT   Region          118..131
FT                   /label= V2
FT                   /note= "Variable region 2"
FT   Region          132..195
FT                   /label= C3
FT                   /note= "Conserved region 3"
FT   Region          196..217
FT                   /label= V3
FT                   /note= "Variable region 3"
FT   Region          218..236
FT                   /label= C4
FT                   /note= "Conserved region 4"
FT   Region          237..243
FT                   /label= V4
FT                   /note= "Variable region 4"
FT   Region          244..599
FT                   /label= C5
FT                   /note= "Conserved region 5"
XX
PN   WO200155182-A1.
XX
PD   02-AUG-2001.
XX
PF   25-JAN-2001; 2001WO-AU00069.
XX
PR   25-JAN-2000; 2000US-0177917.
XX
PA   (UYQU ) UNIV QUEENSLAND.
XX
PI   Peak IRA,  Jennings MP;
XX
DR   WPI; 2001-488774/53.
DR   N-PSDB; AAS09166.
XX
PT   New NhhA surface antigen polypeptides and polynucleotides from
PT   Neisseria meningitidis, useful in producing vaccines for treating or
PT   preventing broad spectrum of Neisseria meningitidis -
XX
PS   Claim 9; Fig 1; 91pp; English.
XX
CC   The present invention relates to the isolation of novel Neisseria
CC   meningitidis mutant polypeptides of the surface antigen NhhA
CC   (AAU06182-AAU06186). The modified or mutant NhhA polypeptides are
CC   characterised by deletions of non-conserved amino acids, particularly
CC   the deletion of variable regions. The deletion mutants are useful in
CC   diagnostics, therapeutic and prophylactic vaccines against a broader
CC   spectrum of N. meningitidis, and in designing and/or screening of
CC   medicaments. The mutant proteins when used as a vaccine can effectively
CC   immunise against a broader spectrum of N. meningitidis strains than
CC   would be expected from a corresponding wild-type surface antigen.
CC   The present sequence representing the wild type surface antigen NhhA
CC   from N. meningitidis strain H38 is 1 of 10 NhhA polypeptide sequences
CC   (AAU06171-AAU06180) from 10 different N. meningitidis strains given in
CC   the present invention.
XX
SQ   Sequence   599 AA;

  Query Match             96.2%;  Score 2260;  DB 22;  Length 599;
  Best Local Similarity   96.5%;  Pred. No. 5.6e-134;
  Matches  445;  Conservative    5;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          2 DETGLINVETEKLSFGANGKKVNIISDTKGLNFAKETAGTNGDTTVHLNGIGSTLTDMLL 61
              | | | :||||||||||| |||| ||||||||||||||||||||||||||||||||| ||
Db        139 DLTDLTSVETEKLSFGANGNKVNITSDTKGLNFAKETAGTNGDTTVHLNGIGSTLTDTLL 198

Qy         62 NTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLS 121
              ||||||||||||||||:|||||||||||||||||||||||||||||||||| |||||||||
Db        199 NTGATTNVTNDNVTDDKKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVHTYDTVEFLS 258

Qy        122 ADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKGKGENGSSTDEGEGLVTAK 181
              ||||||||||||||||:||||||||||||||||||||||||||||||||||||||||||||
Db        259 ADTKTTTVNVESKDNGKRTEVKIGAKTSVIKEKDGKLVTGKGKGENGSSTDEGEGLVTAK 318

Qy        182 EVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTKVTFASGNGTTATVSKDDQGNITVK 241
              |||||||||||||||||||||||||||||||||| |||||| ||||||||||||||||||
Db        319 EVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVK 378

Qy        242 YDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEIT 301
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        379 YDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEIT 438

Qy        302 RNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGALNVGSKDANKPVRITNVAPGVKE 361
              |||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||
Db        439 RNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDKGALNVGSKDANKPVRITNVAPGVKE 498

Qy        362 GDVTNVAQLKGVAQNLNNRIDNVNGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYLG 421
              |||||||||||||||||||||:|||||||||||||||||||||||||||||||||||| |
```

```
Db        499 GDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRG 558

Qy        422 EAGYAIGYSSISAGGNWIIKGTASGNSRGHFGASASVGYQW 462
              ||||||||||| |||||||||||||||||||||||||||||
Db        559 EAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 599


RESULT 10
AAY23739
ID   AAY23739 standard; Protein; 594 AA.
XX
AC   AAY23739;
XX
DT   08-SEP-1999  (first entry)
XX
DE   A surface protein of Neisseria meningitidis.
XX
KW   Surface protein; surface glycoprotein; infection; vaccine;
KW   immunoreactive peptide.
XX
OS   Neisseria meningitidis.
XX
PN   WO9931132-A1.
XX
PD   24-JUN-1999.
XX
PF   14-DEC-1998;   98WO-AU01031.
XX
PR   12-DEC-1997;   97GB-0026398.
XX
PA   (ISIS-) ISIS INNOVATION LTD.
PA   (UYQU ) UNIV QUEENSLAND.
XX
PI   Jennings MP,  Moxon ER,  Peak IRA;
XX
DR   WPI; 1999-418754/35.
DR   N-PSDB; AAX85791.
XX
PT   Neisseria meningitidis surface proteins useful for treating N.
PT   meningitidis infections
XX
PS   Claim 1; Page 95-97; 132pp; English.
XX
CC   The present sequence represents a surface protein of Neiserria
CC   meningitidis which is approximately 62 kDa. The N. meningitidis
CC   surface glycoproteins, nucleic acids, the primers and optionally
CC   a thermostable polymerase, or antibodies are useful in a kit for
CC   the detection or diagnosis of N. meningitidis infection in humans.
CC   The N. meningitidis surface glycoproteins can also be used to
CC   prevent or treat N. meningitidis infection in humans, especially
CC   in the form of vaccines. The proteins and antibodies can also
CC   be used to identify immunoreactive peptides.
XX
SQ   Sequence   594 AA;

  Query Match             95.6%;  Score 2247;  DB 20;  Length 594;
  Best Local Similarity   96.3%;  Pred. No. 3.6e-133;
  Matches  444;  Conservative    2;  Mismatches   15;  Indels    0;  Gaps    0;

Qy          2 DETGLINVETEKLSFGANGKKVNIISDTKGLNFAKETAGTNGDTTVHLNGIGSTLTDMLL 61
              | | |:|||||||||||| |||| |||||||||||||||| ||||||||||||| ||
Db        134 DLTDLTSVETEKLSFGANGNKVNITSDTKGLNFAKETAGTNGDPTVHLNGIGSTLTDTLL 193

Qy         62 NTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLS 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        194 NTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLS 253

Qy        122 ADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKGKGENGSSTDEGEGLVTAK 181
              |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        254 ADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKGKDENGSSTDEGEGLVTAK 313

Qy        182 EVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTKVTFASGNGTTATVSKDDQGNITVK 241
              |||||||||||||||||||||||||||||||| |||||| ||||||||||||||||
Db        314 EVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVK 373

Qy        242 YDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEIT 301
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        374 YDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEIT 433

Qy        302 RNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGALNVGSKDANKPVRITNVAPGVKE 361
              |||||||||||| |||||||||||||||||||||||||||||| ||||||||||||||
Db        434 RNGKNIDIATSMAPQFSSVSLGAGADAPTLSVDDEGALNVGSKDTNKPVRITNVAPGVKE 493

Qy        362 GDVTNVAQLKGVAQNLNNRIDNVNGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYLG 421
              |||||||||||||||||||||:|||||||||||||||||||||||||||| || |
Db        494 GDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGDTYRG 553

Qy        422 EAGYAIGYSSISAGGNWIIKGTASGNSRGHFGASASVGYQW 462
              ||||||||||| |||||||||||||||||||||||||||||
Db        554 EAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 594


RESULT 11
AAU06179
ID   AAU06179 standard; Protein; 594 AA.
XX
AC   AAU06179;
XX
DT   24-OCT-2001  (first entry)
XX
DE   N. meningitidis BZ198 surface antigen NhhA polypeptide sequence.
XX
KW   Surface antigen NhhA; meningococcal disease; meningitis vaccine.
XX
OS   Neisseria meningitidis strain BZ198.
XX
FH   Key             Location/Qualifiers
FT   Region          1..50
FT                   /label= C1
FT                   /note= "Conserved region 1"
FT   Region          51..104
FT                   /label= V1
FT                   /note= "Variable region 1"
FT   Region          105..116
FT                   /label= C2
FT                   /note= "Conserved region 2"
FT   Region          117..126
FT                   /label= V2
FT                   /note= "Variable region 2"
FT   Region          127..190
FT                   /label= C3
FT                   /note= "Conserved region 3"
FT   Region          191..212
FT                   /label= V3
FT                   /note= "Variable region 3"
FT   Region          213..231
FT                   /label= C4
FT                   /note= "Conserved region 4"
FT   Region          232..238
FT                   /label= V4
FT                   /note= "Variable region 4"
FT   Region          239..594
FT                   /label= C5
FT                   /note= "Conserved region 5"
XX
PN   WO200155182-A1.
XX
PD   02-AUG-2001.
XX
PF   25-JAN-2001; 2001WO-AU00069.
XX
PR   25-JAN-2000; 2000US-0177917.
XX
PA   (UYQU ) UNIV QUEENSLAND.
XX
PI   Peak IRA,  Jennings MP;
```

```
XX
DR   WPI; 2001-488774/53.
DR   N-PSDB; AAS09169.
XX
PT   New NhhA surface antigen polypeptides and polynucleotides from
PT   Neisseria meningitidis, useful in producing vaccines for treating or
PT   preventing broad spectrum of Neisseria meningitidis -
XX
PS   Claim 9; Fig 1; 91pp; English.
XX
CC   The present invention relates to the isolation of novel Neisseria
CC   meningitidis mutant polypeptides of the surface antigen NhhA
CC   (AAU06182-AAU06186). The modified or mutant NhhA polypeptides are
CC   characterised by deletions of non-conserved amino acids, particularly
CC   the deletion of variable regions. The deletion mutants are useful in
CC   diagnostics, therapeutic and prophylactic vaccines against a broader
CC   spectrum of N. meningitidis, and in designing and/or screening of
CC   medicaments. The mutant proteins when used as a vaccine can effectively
CC   immunise against a broader spectrum of N. meningitidis strains than
CC   would be expected from a corresponding wild-type surface antigen.
CC   The present sequence representing the wild type surface antigen NhhA
CC   from N. meningitidis strain BZ198 is 1 of 10 NhhA polypeptide sequences
CC   (AAU06171-AAU06180) from 10 different N. meningitidis strains given in
CC   the present invention.
XX
SQ   Sequence   594 AA;

  Query Match             95.6%;  Score 2247;  DB 22;  Length 594;
  Best Local Similarity   96.3%;  Pred. No. 3.6e-133;
  Matches  444;  Conservative    2;  Mismatches   15;  Indels    0;  Gaps    0;

Qy          2 DETGLINVETEKLSFGANGKKVNIISDTKGLNFAKETAGTNGDTTVHLNGIGSTLTDMLL 61
              | | | :|||||||||||| |||| |||||||||||||||||| ||||||||||||| ||
Db        134 DLTDLTSVETEKLSFGANGNKVNITSDTKGLNFAKETAGTNGDPTVHLNGIGSTLTDTLL 193

Qy         62 NTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLS 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        194 NTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLS 253

Qy        122 ADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKGKGENGSSTDEGEGLVTAK 181
              |||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db        254 ADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKGKDENGSSTDEGEGLVTAK 313

Qy        182 EVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTKVTFASGNGTTATVSKDDQGNITVK 241
              |||||||||||||||||||||||||||||||||||||| |||||| ||||||||||||||||
Db        314 EVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVK 373

Qy        242 YDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEIT 301
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        374 YDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEIT 433

Qy        302 RNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGALNVGSKDANKPVRITNVAPGVKE 361
              |||||||||||| |||||||||||||||||||||||||||||| ||||||||||||||||
Db        434 RNGKNIDIATSMAPQFSSVSLGAGADAPTLSVDDEGALNVGSKDTNKPVRITNVAPGVKE 493

Qy        362 GDVTNVAQLKGVAQNLNNRIDNVNGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYLG 421
              |||||||||||||||||||||:|||||||||||||||||||||||||||||| || |
Db        494 GDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGDTYRG 553

Qy        422 EAGYAIGYSSISAGGNWIIKGTASGNSRGHFGASASVGYQW 462
              |||||||||||| |||||||||||||||||||||||||||
Db        554 EAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 594


RESULT 12
AAU06182
ID   AAU06182 standard; Protein; 512 AA.
XX
AC   AAU06182;
XX
DT   24-OCT-2001  (first entry)
XX
DE   N. meningitidis PMC21 NhhA deletion mutant #1.
XX
KW   Surface antigen NhhA; meningococcal disease; meningitis vaccine;
KW   mutant; mutein.
XX
OS   Neisseria meningitidis strain PMC21.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..51
FT                   /label= Signal_peptide
FT   Protein         52..512
FT                   /label= Mature_NhhA_deletion_mutant_#1
FT                   /note= "Predicted mature protein, specifically
FT                         claimed in claim 12"
XX
PN   WO200155182-A1.
XX
PD   02-AUG-2001.
XX
PF   25-JAN-2001; 2001WO-AU00069.
XX
PR   25-JAN-2000; 2000US-0177917.
XX
PA   (UYQU ) UNIV QUEENSLAND.
XX
PI   Peak IRA,  Jennings MP;
XX
DR   WPI; 2001-488774/53.
DR   N-PSDB; AAS09172.
XX
PT   New NhhA surface antigen polypeptides and polynucleotides from
PT   Neisseria meningitidis, useful in producing vaccines for treating or
PT   preventing broad spectrum of Neisseria meningitidis -
XX
PS   Claim 12; Fig 5; 91pp; English.
XX
CC   The present invention relates to the isolation of novel Neisseria
CC   meningitidis mutant polypeptides of the surface antigen NhhA
CC   (AAU06182-AAU06186). The modified or mutant NhhA polypeptides are
CC   characterised by deletions of non-conserved amino acids, particularly
CC   the deletion of variable regions. The deletion mutants are useful in
CC   diagnostics, therapeutic and prophylactic vaccines against a broader
CC   spectrum of N. meningitidis, and in designing and/or screening of
CC   medicaments. The mutant proteins when used as a vaccine can effectively
CC   immunise against a broader spectrum of N. meningitidis strains than
CC   would be expected from a corresponding wild-type surface antigen.
CC   The present sequence represents N. meningitidis strain PMC21 surface
CC   antigen NhhA deletion mutant #1.
XX
SQ   Sequence   512 AA;

  Query Match             94.2%;  Score 2213.5;  DB 22;  Length 512;
  Best Local Similarity   95.4%;  Pred. No. 3.9e-131;
  Matches  440;  Conservative    4;  Mismatches   16;  Indels    1;  Gaps    1;

Qy          2 DETGLINVETEKLSFGANGKKVNIISDTKGLNFAKETAGTNGDTTVHLNGIGSTLTDMLL 61
              :|| | :| |||||| ||| |||| ||||||||||||||||||||||||||||||| ||
Db         53 NETDLTSVGTEKLSFSANGNKVNITSDTKGLNFAKETAGTNGDTTVHLNGIGSTLTDTLL 112

Qy         62 NTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLS 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        113 NTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLS 172

Qy        122 ADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKGKGENGSSTDEGEGLVTAK 181
              |||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db        173 ADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAK 232

Qy        182 EVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTKVTFASGNGTTATVSKDDQGNITVK 241
              |||||||||||||||||||||||||||||||||||||| |||||| ||||||||||||||||
Db        233 EVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVM 292

Qy        242 YDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEIT 301
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        293 YDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEIT 352

Qy        302 RNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGALNVGSKDANKPVRITNVAPGVKE 361
              |||||||||||||||||||||||||||||||| : |||||||  ||||||||||||||||
Db        353 RNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGD-ALNVGSKKDNKPVRITNVAPGVKE 411

Qy        362 GDVTNVAQLKGVAQNLNNRIDNVNGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYLG 421
              |||||||||||||||||||||:|||||||||||||||||||||||||||||||||||||| |
Db        412 GDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRG 471

Qy        422 EAGYAIGYSSISAGGNWIIKGTASGNSRGHFGASASVGYQW 462
              |||||||||||| |||||||||||||||||||||||||||
Db        472 EAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 512


RESULT 13
AAY23740
ID   AAY23740 standard; Protein; 594 AA.
XX
AC   AAY23740;
XX
DT   08-SEP-1999  (first entry)
XX
DE   A surface protein of Neisseria meningitidis.
XX
KW   Surface protein; surface glycoprotein; infection; vaccine;
KW   immunoreactive peptide.
XX
OS   Neisseria meningitidis.
XX
PN   WO9931132-A1.
XX
PD   24-JUN-1999.
XX
PF   14-DEC-1998;   98WO-AU01031.
XX
PR   12-DEC-1997;   97GB-0026398.
XX
PA   (ISIS-) ISIS INNOVATION LTD.
PA   (UYQU ) UNIV QUEENSLAND.
XX
PI   Jennings MP,  Moxon ER,  Peak IRA;
XX
DR   WPI; 1999-418754/35.
DR   N-PSDB; AAX85792.
XX
PT   Neisseria meningitidis surface proteins useful for treating N.
PT   meningitidis infections
XX
PS   Claim 1; Page 100-101; 132pp; English.
XX
CC   The present sequence represents a surface protein of Neiserria
CC   meningitidis which is approximately 62 kDa. The N. meningitidis
CC   surface glycoproteins, nucleic acids, the primers and optionally
CC   a thermostable polymerase, or antibodies are useful in a kit for
CC   the detection or diagnosis of N. meningitidis infection in humans.
CC   The N. meningitidis surface glycoproteins can also be used to
CC   prevent or treat N. meningitidis infection in humans, especially
CC   in the form of vaccines. The proteins and antibodies can also
CC   be used to identify immunoreactive peptides.
XX
SQ   Sequence   594 AA;

  Query Match            94.2%;  Score 2213;  DB 20;  Length 594;
  Best Local Similarity  95.0%;  Pred. No. 5e-131;
  Matches  438;  Conservative    5;  Mismatches   18;  Indels    0;  Gaps    0;

Qy          2 DETGLINVETEKLSFGANGKKVNIISDTKGLNFAKETAGTNGDTTVHLNGIGSTLTDMLL 61
              | | | :| |||||| ||  |||| ||||||||||:|| ||||||||||||||||| ||
Db        134 DLTDLTSVGTEKLSFSANSNKVNITSDTKGLNFAKKTAETNGDTTVHLNGIGSTLTDTLL 193

Qy         62 NTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLS 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        194 NTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLS 253

Qy        122 ADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKGKGENGSSTDEGEGLVTAK 181
              |||||||||||||||||:|||||||||||||||||||| |||| ||||:||||||||
Db        254 ADTKTTTVNVESKDNGKRTEVKIGAKTSVIKEKDGKLVTGKDKGENDSSTDKGEGLVTAK 313

Qy        182 EVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTKVTFASGNGTTATVSKDDQGNITVK 241
              |||||||||||||||||||||||||||||||||||| |||||| ||||||||||||||||
Db        314 EVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVM 373

Qy        242 YDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEIT 301
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        374 YDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEIT 433

Qy        302 RNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGALNVGSKDANKPVRITNVAPGVKE 361
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        434 RNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGALNVGSKDANKPVRITNVAPGVKE 493

Qy        362 GDVTNVAQLKGVAQNLNNRIDNVNGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYLG 421
              |||||||||||||||| ||||:|||||||||||||||||||||||||||||||||||| |
Db        494 GDVTNVAQLKGVAQNLNNHIDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRG 553

Qy        422 EAGYAIGYSSISAGGNWIIKGTASGNSRGHFGASASVGYQW 462
              |||||||||||| |||||||||||||||||||||||||||
Db        554 EAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 594


RESULT 14
AAY57044
ID   AAY57044 standard; Protein; 594 AA.
XX
AC   AAY57044;
XX
DT   21-FEB-2000  (first entry)
XX
DE   BASB029 amino acid sequence from N. meningitidis strain ATCC13090.
XX
KW   BASB029; Nisseria meningitidis; surface fibril protein; HSF; diagnosis;
KW   infection; treatment; prevent; antibacterial drug.
XX
OS   Neisseria meningitidis.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 104
FT                   /note= "Encoded by AATC"
XX
PN   WO9958683-A2.
XX
PD   18-NOV-1999.
XX
PF   07-MAY-1999;   99WO-EP03255.
XX
PR   13-MAY-1998;   98GB-0010276.
XX
PA   (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PI   Ruelle J;
XX
DR   WPI; 2000-053103/04.
DR   N-PSDB; AAZ39864.
XX
PT   New polypeptide from neisseria meningitidis useful for diagnosis,
PT   treatment or prevention of bacterial infections in mammal  -
XX
PS   Claim 4; Fig 2; 74pp; English.
XX
CC   This is the Nisseria meningitidis BASB029 amino acid sequence from
CC   serogroup B strain ATCC13090. The BASB029 protein is homologous to the
CC   Haemophilus influenzae surface fibril (HSF) protein. The invention
CC   relates to BASB029 polynucleotide sequences (AAZ39864-Z39865) and
CC   polypeptide sequences (AAY57044-Y57045) and their immunogenic fragments.
CC   BASB029 polypeptides are useful in a method of diagnosing a Neisseria
CC   meningitidis infection in a mammal. Compositions containing BASB029
```

```
CC   polynucleotides and polypeptides are useful for generating an immune
CC   response in an animal. A therapeutic composition comprising an antibody
CC   directed against BAsB029 is useful in treating humans with Neisseria
CC   meningitidis disease. The polynucleotide is useful in the diagnosis of
CC   the stage of infection, type of infection, susceptibility to an
CC   infection which results from increased or decreased expression of the
CC   polynucleotide, and for therapeutic or prophylactic purposes,
CC   particularly genetic immunisation. Antibodies against BASB029
CC   polynucleotides and polypeptides are also useful for treating infections
CC   particularly bacterial infections. The protein is useful in the
CC   screening and development of antibacterial drugs. Fused recombinant
CC   protein is useful for the stimulation of the immune system of an organism
CC   receiving the protein.
XX
SQ   Sequence   594 AA;

  Query Match             94.2%;  Score 2213;  DB 21;  Length 594;
  Best Local Similarity   95.0%;  Pred. No. 5e-131;
  Matches  438;  Conservative    5;  Mismatches   18;  Indels    0;  Gaps    0;

Qy          2 DETGLINVETEKLSFGANGKKVNIISDTKGLNFAKETAGTNGDTTVHLNGIGSTLTDMLL 61
              | | | :| |||||| ||  |||| |||||||||:|| |||||||||||||||||| ||
Db        134 DLTDLTSVGTEKLSFSANSNKVNITSDTKGLNFAKKTAETNGDTTVHLNGIGSTLTDTLL 193

Qy         62 NTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLS 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        194 NTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLS 253

Qy        122 ADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKGKGENGSSTDEGEGLVTAK 181
              ||||||||||||||||:|||||||||||||||||||| |||| ||||:||||||||
Db        254 ADTKTTTVNVESKDNGKRTEVKIGAKTSVIKEKDGKLVTGKDKGENDSSTDKGEGLVTAK 313

Qy        182 EVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTKVTFASGNGTTATVSKDDQGNITVK 241
              |||||||||||||||||||||||||||||||| |||||| ||||||||||||||||
Db        314 EVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVM 373

Qy        242 YDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEIT 301
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        374 YDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEIT 433

Qy        302 RNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGALNVGSKDANKPVRITNVAPGVKE 361
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        434 RNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGALNVGSKDANKPVRITNVAPGVKE 493

Qy        362 GDVTNVAQLKGVAQNLNNRIDNVNGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYLG 421
              |||||||||||||||| ||||:|||||||||||||||||||||||||||||||||| |
Db        494 GDVTNVAQLKGVAQNLNNHIDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRG 553

Qy        422 EAGYAIGYSSISAGGNWIIKGTASGNSRGHFGASASVGYQW 462
              |||||||||||| ||||||||||||||||||||||||||
Db        554 EAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 594


RESULT 15
AAU06174
ID   AAU06174 standard; Protein; 594 AA.
XX
AC   AAU06174;
XX
DT   24-OCT-2001  (first entry)
XX
DE   N. meningitidis EG327 surface antigen NhhA polypeptide sequence.
XX
KW   Surface antigen NhhA; meningococcal disease; meningitis vaccine.
XX
OS   Neisseria meningitidis strain EG327.
XX
FH   Key             Location/Qualifiers
FT   Region          1..50
FT                   /label= C1
FT                   /note= "Conserved region 1"
FT   Region          51..104
FT                   /label= V1
FT                   /note= "Variable region 1"
FT   Region          105..116
FT                   /label= C2
FT                   /note= "Conserved region 2"
FT   Region          117..126
FT                   /label= V2
FT                   /note= "Variable region 2"
FT   Region          127..190
FT                   /label= C3
FT                   /note= "Conserved region 3"
FT   Region          191..212
FT                   /label= V3
FT                   /note= "Variable region 3"
FT   Region          213..231
FT                   /label= C4
FT                   /note= "Conserved region 4"
FT   Region          232..238
FT                   /label= V4
FT                   /note= "Variable region 4"
FT   Region          239..594
FT                   /label= C5
FT                   /note= "Conserved region 5"
XX
PN   WO200155182-A1.
XX
PD   02-AUG-2001.
XX
PF   25-JAN-2001; 2001WO-AU00069.
XX
PR   25-JAN-2000; 2000US-0177917.
XX
PA   (UYQU ) UNIV QUEENSLAND.
XX
PI   Peak IRA,  Jennings MP;
XX
DR   WPI; 2001-488774/53.
DR   N-PSDB; AAS09164.
XX
PT   New NhhA surface antigen polypeptides and polynucleotides from
PT   Neisseria meningitidis, useful in producing vaccines for treating or
PT   preventing broad spectrum of Neisseria meningitidis -
XX
PS   Claim 9; Fig 1; 91pp; English.
XX
CC   The present invention relates to the isolation of novel Neisseria
CC   meningitidis mutant polypeptides of the surface antigen NhhA
CC   (AAU06182-AAU06186). The modified or mutant NhhA polypeptides are
CC   characterised by deletions of non-conserved amino acids, particularly
CC   the deletion of variable regions. The deletion mutants are useful in
CC   diagnostics, therapeutic and prophylactic vaccines against a broader
CC   spectrum of N. meningitidis, and in designing and/or screening of
CC   medicaments. The mutant proteins when used as a vaccine can effectively
CC   immunise against a broader spectrum of N. meningitidis strains than
CC   would be expected from a corresponding wild-type surface antigen.
CC   The present sequence representing the wild type surface antigen NhhA
CC   from N. meningitidis strain EG327 is 1 of 10 NhhA polypeptide sequences
CC   (AAU06171-AAU06180) from 10 different N. meningitidis strains given in
CC   the present invention.
XX
SQ   Sequence   594 AA;

  Query Match             94.2%;  Score 2213;  DB 22;  Length 594;
  Best Local Similarity   95.0%;  Pred. No. 5e-131;
  Matches  438;  Conservative    5;  Mismatches   18;  Indels    0;  Gaps    0;

Qy          2 DETGLINVETEKLSFGANGKKVNIISDTKGLNFAKETAGTNGDTTVHLNGIGSTLTDMLL 61
              | | | :| |||||| ||  |||| |||||||||:|| |||||||||||||||||| ||
Db        134 DLTDLTSVGTEKLSFSANSNKVNITSDTKGLNFAKKTAETNGDTTVHLNGIGSTLTDTLL 193

Qy         62 NTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLS 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        194 NTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLS 253
```

```
Qy        122 ADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKGKGENGSSTDEGEGLVTAK 181
              ||||||||||||||||||:||||||||||||||||||||| |||| |||||:||||||||
Db        254 ADTKTTTVNVESKDNGKRTEVKIGAKTSVIKEKDGKLVTGKDKGENDSSTDKGEGLVTAK 313

Qy        182 EVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTKVTFASGNGTTATVSKDDQGNITVK 241
              |||||||||||||||||||||||||||||||||||| |||||| ||||||||||||||||
Db        314 EVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVM 373

Qy        242 YDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEIT 301
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        374 YDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEIT 433

Qy        302 RNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGALNVGSKDANKPVRITNVAPGVKE 361
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        434 RNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGALNVGSKDANKPVRITNVAPGVKE 493

Qy        362 GDVTNVAQLKGVAQNLNNRIDNVNGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYLG 421
              |||||||||||||||| ||||:|||||||||||||||||||||||||||||||||| |
Db        494 GDVTNVAQLKGVAQNLNNHIDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRG 553

Qy        422 EAGYAIGYSSISAGGNWIIKGTASGNSRGHFGASASVGYQW 462
              |||||||||||| ||||||||||||||||||||||||||||
Db        554 EAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 594
```

Search completed: October 6, 2003, 09:22:42
Job time : 40.6746 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: October 6, 2003, 09:14:45 ; Search time 12.3548 Seconds
(without alignments)
1582.188 Million cell updates/sec

Title: US-09-771-382-36
Perfect score: 2350
Sequence: 1 TDETGLINVETEKLSFGANG.....TASGNSRGHFGASASVGYQW 462

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2337	99.4	592	3	US-09-377-155-17	Sequence 17, Appl
2	2337	99.4	592	4	US-09-669-974-17	Sequence 17, Appl
3	2261	96.2	598	3	US-09-377-155-5	Sequence 5, Appli
4	2261	96.2	598	3	US-09-377-155-13	Sequence 13, Appl
5	2261	96.2	598	4	US-09-669-974-5	Sequence 5, Appli
6	2261	96.2	598	4	US-09-669-974-13	Sequence 13, Appl
7	2260	96.2	599	3	US-09-377-155-15	Sequence 15, Appl
8	2260	96.2	599	4	US-09-669-974-15	Sequence 15, Appl
9	2247	95.6	594	3	US-09-377-155-7	Sequence 7, Appli
10	2247	95.6	594	4	US-09-669-974-7	Sequence 7, Appli
11	2213	94.2	594	3	US-09-377-155-9	Sequence 9, Appli
12	2213	94.2	594	4	US-09-669-974-9	Sequence 9, Appli
13	2210.5	94.1	591	3	US-09-377-155-11	Sequence 11, Appl
14	2210.5	94.1	591	3	US-09-377-155-21	Sequence 21, Appl
15	2210.5	94.1	591	4	US-09-669-974-11	Sequence 11, Appl
16	2210.5	94.1	591	4	US-09-669-974-21	Sequence 21, Appl
17	2206.5	93.9	592	3	US-09-377-155-2	Sequence 2, Appli
18	2206.5	93.9	592	4	US-09-669-974-2	Sequence 2, Appli
19	2116	90.0	589	3	US-09-377-155-19	Sequence 19, Appl
20	2116	90.0	589	4	US-09-669-974-19	Sequence 19, Appl
21	876	37.3	2353	3	US-09-377-155-33	Sequence 33, Appl
22	876	37.3	2353	3	US-08-913-942-4	Sequence 4, Appli
23	876	37.3	2353	4	US-09-669-974-33	Sequence 33, Appl
24	876	37.3	2411	4	US-09-268-347-36	Sequence 36, Appl
25	869.5	37.0	2354	4	US-09-268-347-47	Sequence 47, Appl
26	865	36.8	1094	4	US-09-268-347-32	Sequence 32, Appl
27	855	36.4	1912	1	US-08-409-995-4	Sequence 4, Appli
28	855	36.4	1912	3	US-08-685-467-4	Sequence 4, Appli
29	849.5	36.1	1098	1	US-08-409-995-2	Sequence 2, Appli
30	849.5	36.1	1098	3	US-08-685-467-2	Sequence 2, Appli
31	849.5	36.1	1098	3	US-09-377-155-32	Sequence 32, Appl
32	849.5	36.1	1098	3	US-08-913-942-2	Sequence 2, Appli
33	849.5	36.1	1098	4	US-09-669-974-32	Sequence 32, Appl
34	849.5	36.1	1098	4	US-09-268-347-44	Sequence 44, Appl
35	846.5	36.0	658	1	US-08-409-995-5	Sequence 5, Appli
36	846.5	36.0	658	3	US-08-685-467-5	Sequence 5, Appli
37	846.5	36.0	658	3	US-08-913-942-5	Sequence 5, Appli
38	767.5	32.7	607	1	US-08-409-995-6	Sequence 6, Appli
39	767.5	32.7	607	3	US-08-685-467-6	Sequence 6, Appli
40	767.5	32.7	607	3	US-08-913-942-6	Sequence 6, Appli
41	613	26.1	679	3	US-08-913-942-15	Sequence 15, Appl
42	613	26.1	679	4	US-09-268-347-26	Sequence 26, Appl
43	571.5	24.3	1002	4	US-09-268-347-24	Sequence 24, Appl
44	551.5	23.5	1004	4	US-09-268-347-30	Sequence 30, Appl
45	511	21.7	1104	4	US-09-268-347-28	Sequence 28, Appl

ALIGNMENTS

RESULT 1
US-09-377-155-17
; Sequence 17, Application US/09377155
; Patent No. 6197312
; GENERAL INFORMATION:
; APPLICANT: PEAK, Ian Richard Anselm
; APPLICANT: JENNINGS, Michael Paul
; APPLICANT: MOXON, E. Richard
; TITLE OF INVENTION: NOVEL SURFACE ANTIGEN
; FILE REFERENCE: 065064/0128
; CURRENT APPLICATION NUMBER: US/09/377,155
; CURRENT FILING DATE: 1999-08-19
; PRIOR APPLICATION NUMBER: PCT/AU98/01031
; PRIOR FILING DATE: 1998-12-14
; PRIOR APPLICATION NUMBER: GB 9726398.2
; PRIOR FILING DATE: 1997-12-12
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17
; LENGTH: 592
; TYPE: PRT
; ORGANISM: Neisseria meningitidis
US-09-377-155-17

Query Match 99.4%; Score 2337; DB 3; Length 592;
Best Local Similarity 99.8%; Pred. No. 2.7e-178;
Matches 460; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          2 DETGLINVETEKLSFGANGKKVNIISDTKGLNFAKETAGTNGDTTVHLNGIGSTLTDMLL 61
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        132 DLTGLINVETEKLSFGANGKKVNIISDTKGLNFAKETAGTNGDTTVHLNGIGSTLTDMLL 191

Qy         62 NTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLS 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        192 NTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLS 251

Qy        122 ADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKGKGENGSSTDEGEGLVTAK 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        252 ADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKGKGENGSSTDEGEGLVTAK 311

Qy        182 EVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTKVTFASGNGTTATVSKDDQGNITVK 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        312 EVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTKVTFASGNGTTATVSKDDQGNITVK 371

Qy        242 YDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEIT 301
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        372 YDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEIT 431

Qy        302 RNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGALNVGSKDANKPVRITNVAPGVKE 361
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        432 RNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGALNVGSKDANKPVRITNVAPGVKE 491

Qy        362 GDVTNVAQLKGVAQNLNNRIDNVNGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYLG 421
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        492 GDVTNVAQLKGVAQNLNNRIDNVNGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYLG 551

Qy        422 EAGYAIGYSSISAGGNWIIKGTASGNSRGHFGASASVGYQW 462
              ||||||||||||||||||||||||||||||||||||||||
Db        552 EAGYAIGYSSISAGGNWIIKGTASGNSRGHFGASASVGYQW 592


RESULT 2
US-09-669-974-17
; Sequence 17, Application US/09669974
; Patent No. 6333173
; GENERAL INFORMATION:
;  APPLICANT: PEAK, Ian Richard Anselm
;  APPLICANT:  JENNINGS, Michael Paul
;  APPLICANT:  MOXON, E. Richard
;  TITLE OF INVENTION: NOVEL SURFACE ANTIGEN
;  FILE REFERENCE: 065064/0128
;  CURRENT APPLICATION NUMBER: US/09/669,974
;  CURRENT FILING DATE:  2000-09-26
;  PRIOR APPLICATION NUMBER: US 09/377,155
;  PRIOR FILING DATE: 1999-08-19
;  PRIOR APPLICATION NUMBER: PCT/AU98/01031
;  PRIOR FILING DATE: 1998-12-14
;  PRIOR APPLICATION NUMBER: GB 9726398.2
;  PRIOR FILING DATE: 1997-12-12
;  NUMBER OF SEQ ID NOS: 33
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17
;   LENGTH: 592
;   TYPE: PRT
;   ORGANISM: Neisseria meningitidis
US-09-669-974-17

  Query Match             99.4%;  Score 2337;  DB 4;  Length 592;
  Best Local Similarity   99.8%;  Pred. No. 2.7e-178;
  Matches  460;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 DETGLINVETEKLSFGANGKKVNIISDTKGLNFAKETAGTNGDTTVHLNGIGSTLTDMLL 61
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        132 DLTGLINVETEKLSFGANGKKVNIISDTKGLNFAKETAGTNGDTTVHLNGIGSTLTDMLL 191

Qy         62 NTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLS 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        192 NTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLS 251

Qy        122 ADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKGKGENGSSTDEGEGLVTAK 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        252 ADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKGKGENGSSTDEGEGLVTAK 311

Qy        182 EVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTKVTFASGNGTTATVSKDDQGNITVK 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        312 EVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTKVTFASGNGTTATVSKDDQGNITVK 371

Qy        242 YDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEIT 301
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        372 YDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEIT 431

Qy        302 RNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGALNVGSKDANKPVRITNVAPGVKE 361
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        432 RNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGALNVGSKDANKPVRITNVAPGVKE 491

Qy        362 GDVTNVAQLKGVAQNLNNRIDNVNGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYLG 421
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        492 GDVTNVAQLKGVAQNLNNRIDNVNGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYLG 551

Qy        422 EAGYAIGYSSISAGGNWIIKGTASGNSRGHFGASASVGYQW 462
              ||||||||||||||||||||||||||||||||||||||||
Db        552 EAGYAIGYSSISAGGNWIIKGTASGNSRGHFGASASVGYQW 592


RESULT 3
US-09-377-155-5
; Sequence 5, Application US/09377155
; Patent No. 6197312
; GENERAL INFORMATION:
;  APPLICANT:  PEAK, Ian Richard Anselm
;  APPLICANT:  JENNINGS, Michael Paul
;  APPLICANT:  MOXON, E. Richard
;  TITLE OF INVENTION:  NOVEL SURFACE ANTIGEN
;  FILE REFERENCE:  065064/0128
;  CURRENT APPLICATION NUMBER: US/09/377,155
;  CURRENT FILING DATE:  1999-08-19
;  PRIOR APPLICATION NUMBER:  PCT/AU98/01031
;  PRIOR FILING DATE:  1998-12-14
;  PRIOR APPLICATION NUMBER:  GB 9726398.2
;  PRIOR FILING DATE:  1997-12-12
;  NUMBER OF SEQ ID NOS:  33
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 5
;   LENGTH: 598
;   TYPE: PRT
;   ORGANISM: Neisseria meningitidis
US-09-377-155-5

  Query Match             96.2%;  Score 2261;  DB 3;  Length 598;
  Best Local Similarity   96.5%;  Pred. No. 3.1e-172;
  Matches  445;  Conservative    4;  Mismatches   12;  Indels    0;  Gaps    0;

Qy          2 DETGLINVETEKLSFGANGKKVNIISDTKGLNFAKETAGTNGDTTVHLNGIGSTLTDMLL 61
              | | | :||||||||||||| |||| |||||||||||||||||| |||||||||||| ||
Db        138 DLTDLTSVETEKLSFGANGNKVNITSDTKGLNFAKETAGTNGDPTVHLNGIGSTLTDTLL 197

Qy         62 NTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLS 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        198 NTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLS 257

Qy        122 ADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKGKGENGSSTDEGEGLVTAK 181
              |||||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||
Db        258 ADTKTTTVNVESKDNGKRTEVKIGAKTSVIKEKDGKLVTGKGKGENGSSTDEGEGLVTAK 317

Qy        182 EVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTKVTFASGNGTTATVSKDDQGNITVK 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        318 EVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTKVTFASGNGTTATVSKDDQGNITVK 377

Qy        242 YDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEIT 301
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        378 YDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEIT 437

Qy        302 RNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGALNVGSKDANKPVRITNVAPGVKE 361
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        438 RNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGALNVGSKDANKPVRITNVAPGVKE 497

Qy        362 GDVTNVAQLKGVAQNLNNRIDNVNGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYLG 421
              ||||||||||||||||||||||:|||||||||||||||| |||||||||||||||| |
Db        498 GDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLAQAYLPGKSMMAIGGGTYRG 557

Qy        422 EAGYAIGYSSISAGGNWIIKGTASGNSRGHFGASASVGYQW 462
              |||||||||||  |||:|||||||||||||| ||||||||
Db        558 EAGYAIGYSSISDTGNWVIKGTASGNSRGHFGTSASVGYQW 598


RESULT 4
US-09-377-155-13
; Sequence 13, Application US/09377155
; Patent No. 6197312
; GENERAL INFORMATION:
;  APPLICANT:  PEAK, Ian Richard Anselm
;  APPLICANT:  JENNINGS, Michael Paul
;  APPLICANT:  MOXON, E. Richard
;  TITLE OF INVENTION:  NOVEL SURFACE ANTIGEN
;  FILE REFERENCE:  065064/0128
;  CURRENT APPLICATION NUMBER: US/09/377,155
```

```
;  CURRENT FILING DATE:  1999-08-19
;  PRIOR APPLICATION NUMBER:  PCT/AU98/01031
;  PRIOR FILING DATE:  1998-12-14
;  PRIOR APPLICATION NUMBER:  GB 9726398.2
;  PRIOR FILING DATE:  1997-12-12
;  NUMBER OF SEQ ID NOS:  33
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 13
;   LENGTH: 598
;   TYPE: PRT
;   ORGANISM: Neisseria meningitidis
US-09-377-155-13

  Query Match             96.2%;  Score 2261;  DB 3;  Length 598;
  Best Local Similarity   96.7%;  Pred. No. 3.1e-172;
  Matches  446;  Conservative    3;  Mismatches   12;  Indels    0;  Gaps    0;

Qy          2 DETGLINVETEKLSFGANGKKVNIISDTKGLNFAKETAGTNGDTTVHLNGIGSTLTDMLL 61
              | | | :||||||||||| |||| |||||||||||||||| |||||||||||| ||
Db        138 DLTDLTSVETEKLSFGANGNKVNITSDTKGLNFAKETAGTNGDPTVHLNGIGSTLTDTLL 197

Qy         62 NTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLS 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        198 NTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLS 257

Qy        122 ADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKGKGENGSSTDEGEGLVTAK 181
              |||||||||||||||||||||||||||||||||| |||||||||||||||
Db        258 ADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKGKDENGSSTDEGEGLVTAK 317

Qy        182 EVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTKVTFASGNGTTATVSKDDQGNITVK 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        318 EVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTKVTFASGNGTTATVSKDDQGNITVK 377

Qy        242 YDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEIT 301
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        378 YDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEIT 437

Qy        302 RNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGALNVGSKDANKPVRITNVAPGVKE 361
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        438 RNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGALNVGSKDANKPVRITNVAPGVKE 497

Qy        362 GDVTNVAQLKGVAQNLNNRIDNVNGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYLG 421
              ||||||||||||||||||||:|||||||||||||| |||||||||||||| |
Db        498 GDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLAQAYLPGKSMMAIGGGTYRG 557

Qy        422 EAGYAIGYSSISAGGNWIIKGTASGNSRGHFGASASVGYQW 462
              ||||||||||||  |||:||||||||||||||||||||||
Db        558 EAGYAIGYSSISDTGNWVIKGTASGNSRGHFGASASVGYQW 598


RESULT 5
US-09-669-974-5
; Sequence 5, Application US/09669974
; Patent No. 6333173
; GENERAL INFORMATION:
;  APPLICANT: PEAK, Ian Richard Anselm
;  APPLICANT:  JENNINGS, Michael Paul
;  APPLICANT:  MOXON, E. Richard
;  TITLE OF INVENTION: NOVEL SURFACE ANTIGEN
;  FILE REFERENCE: 065064/0128
;  CURRENT APPLICATION NUMBER: US/09/669,974
;  CURRENT FILING DATE:  2000-09-26
;  PRIOR APPLICATION NUMBER: US 09/377,155
;  PRIOR FILING DATE: 1999-08-19
;  PRIOR APPLICATION NUMBER: PCT/AU98/01031
;  PRIOR FILING DATE: 1998-12-14
;  PRIOR APPLICATION NUMBER: GB 9726398.2
;  PRIOR FILING DATE: 1997-12-12
;  NUMBER OF SEQ ID NOS: 33
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
;   LENGTH: 598
;   TYPE: PRT
;   ORGANISM: Neisseria meningitidis
US-09-669-974-5

  Query Match             96.2%;  Score 2261;  DB 4;  Length 598;
  Best Local Similarity   96.5%;  Pred. No. 3.1e-172;
  Matches  445;  Conservative    4;  Mismatches   12;  Indels    0;  Gaps    0;

Qy          2 DETGLINVETEKLSFGANGKKVNIISDTKGLNFAKETAGTNGDTTVHLNGIGSTLTDMLL 61
              | | | :||||||||||| |||| |||||||||||||||| |||||||||||| ||
Db        138 DLTDLTSVETEKLSFGANGNKVNITSDTKGLNFAKETAGTNGDPTVHLNGIGSTLTDTLL 197

Qy         62 NTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLS 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        198 NTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLS 257

Qy        122 ADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKGKGENGSSTDEGEGLVTAK 181
              |||||||||||||||||||:||||||||||||||||||||||||||||||||||||||||
Db        258 ADTKTTTVNVESKDNGKRTEVKIGAKTSVIKEKDGKLVTGKGKGENGSSTDEGEGLVTAK 317

Qy        182 EVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTKVTFASGNGTTATVSKDDQGNITVK 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        318 EVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTKVTFASGNGTTATVSKDDQGNITVK 377

Qy        242 YDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEIT 301
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        378 YDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEIT 437

Qy        302 RNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGALNVGSKDANKPVRITNVAPGVKE 361
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        438 RNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGALNVGSKDANKPVRITNVAPGVKE 497

Qy        362 GDVTNVAQLKGVAQNLNNRIDNVNGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYLG 421
              ||||||||||||||||||||:|||||||||||||| |||||||||||||| |
Db        498 GDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLAQAYLPGKSMMAIGGGTYRG 557

Qy        422 EAGYAIGYSSISAGGNWIIKGTASGNSRGHFGASASVGYQW 462
              ||||||||||||  |||:||||||||||||||||||||||
Db        558 EAGYAIGYSSISDTGNWVIKGTASGNSRGHFGTSASVGYQW 598


RESULT 6
US-09-669-974-13
; Sequence 13, Application US/09669974
; Patent No. 6333173
; GENERAL INFORMATION:
;  APPLICANT: PEAK, Ian Richard Anselm
;  APPLICANT:  JENNINGS, Michael Paul
;  APPLICANT:  MOXON, E. Richard
;  TITLE OF INVENTION: NOVEL SURFACE ANTIGEN
;  FILE REFERENCE: 065064/0128
;  CURRENT APPLICATION NUMBER: US/09/669,974
;  CURRENT FILING DATE:  2000-09-26
;  PRIOR APPLICATION NUMBER: US 09/377,155
;  PRIOR FILING DATE: 1999-08-19
;  PRIOR APPLICATION NUMBER: PCT/AU98/01031
;  PRIOR FILING DATE: 1998-12-14
;  PRIOR APPLICATION NUMBER: GB 9726398.2
;  PRIOR FILING DATE: 1997-12-12
;  NUMBER OF SEQ ID NOS: 33
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
;   LENGTH: 598
;   TYPE: PRT
;   ORGANISM: Neisseria meningitidis
US-09-669-974-13

  Query Match             96.2%;  Score 2261;  DB 4;  Length 598;
  Best Local Similarity   96.7%;  Pred. No. 3.1e-172;
  Matches  446;  Conservative    3;  Mismatches   12;  Indels    0;  Gaps    0;

Qy          2 DETGLINVETEKLSFGANGKKVNIISDTKGLNFAKETAGTNGDTTVHLNGIGSTLTDMLL 61
              | | | :||||||||||| |||| |||||||||||||||| |||||||||||| ||
Db        138 DLTDLTSVETEKLSFGANGNKVNITSDTKGLNFAKETAGTNGDPTVHLNGIGSTLTDTLL 197
```

```
Qy         62 NTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLS 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        198 NTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLS 257

Qy        122 ADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKGKGENGSSTDEGEGLVTAK 181
              |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        258 ADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKGKDENGSSTDEGEGLVTAK 317

Qy        182 EVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTKVTFASGNGTTATVSKDDQGNITVK 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        318 EVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTKVTFASGNGTTATVSKDDQGNITVK 377

Qy        242 YDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEIT 301
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        378 YDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEIT 437

Qy        302 RNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGALNVGSKDANKPVRITNVAPGVKE 361
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        438 RNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGALNVGSKDANKPVRITNVAPGVKE 497

Qy        362 GDVTNVAQLKGVAQNLNNRIDNVNGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYLG 421
              |||||||||||||||||||||:|||||||||||||||| |||||||||||||||||| |
Db        498 GDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLAQAYLPGKSMMAIGGGTYRG 557

Qy        422 EAGYAIGYSSISAGGNWIIKGTASGNSRGHFGASASVGYQW 462
              ||||||||||||  |||:|||||||||||||||||||||||
Db        558 EAGYAIGYSSISDTGNWVIKGTASGNSRGHFGASASVGYQW 598
```

RESULT 7
US-09-377-155-15
; Sequence 15, Application US/09377155
; Patent No. 6197312
; GENERAL INFORMATION:
; APPLICANT: PEAK, Ian Richard Anselm
; APPLICANT: JENNINGS, Michael Paul
; APPLICANT: MOXON, E. Richard
; TITLE OF INVENTION: NOVEL SURFACE ANTIGEN
; FILE REFERENCE: 065064/0128
; CURRENT APPLICATION NUMBER: US/09/377,155
; CURRENT FILING DATE: 1999-08-19
; PRIOR APPLICATION NUMBER: PCT/AU98/01031
; PRIOR FILING DATE: 1998-12-14
; PRIOR APPLICATION NUMBER: GB 9726398.2
; PRIOR FILING DATE: 1997-12-12
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
; LENGTH: 599
; TYPE: PRT
; ORGANISM: Neisseria meningitidis
US-09-377-155-15

Query Match 96.2%; Score 2260; DB 3; Length 599;
Best Local Similarity 96.5%; Pred. No. 3.8e-172;
Matches 445; Conservative 5; Mismatches 11; Indels 0; Gaps 0;

```
Qy          2 DETGLINVETEKLSFGANGKKVNIISDTKGLNFAKETAGTNGDTTVHLNGIGSTLTDMLL 61
              | | | :||||||||||||| |||| ||||||||||||||||||||||||||||||| ||
Db        139 DLTDLTSVETEKLSFGANGNKVNITSDTKGLNFAKETAGTNGDTTVHLNGIGSTLTDTLL 198

Qy         62 NTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLS 121
              ||||||||||||||:|||||||||||||||||||||||||||||| |||||||||
Db        199 NTGATTNVTNDNVTDDKKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVHTYDTVEFLS 258

Qy        122 ADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKGKGENGSSTDEGEGLVTAK 181
              |||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||
Db        259 ADTKTTTVNVESKDNGKRTEVKIGAKTSVIKEKDGKLVTGKGKGENGSSTDEGEGLVTAK 318

Qy        182 EVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTKVTFASGNGTTATVSKDDQGNITVK 241
              |||||||||||||||||||||||||||||||| |||||| ||||||||||||||||||||
Db        319 EVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVK 378

Qy        242 YDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEIT 301
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        379 YDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEIT 438

Qy        302 RNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGALNVGSKDANKPVRITNVAPGVKE 361
              |||||||||||||||||||||||||||||||||:||||||||||||||||||||||||||
Db        439 RNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDKGALNVGSKDANKPVRITNVAPGVKE 498

Qy        362 GDVTNVAQLKGVAQNLNNRIDNVNGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYLG 421
              |||||||||||||||||||||:|||||||||||||||||||||||||||||||||||| |
Db        499 GDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRG 558

Qy        422 EAGYAIGYSSISAGGNWIIKGTASGNSRGHFGASASVGYQW 462
              |||||||||||| ||||||||||||||||||||||||||||
Db        559 EAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 599
```

RESULT 8
US-09-669-974-15
; Sequence 15, Application US/09669974
; Patent No. 6333173
; GENERAL INFORMATION:
; APPLICANT: PEAK, Ian Richard Anselm
; APPLICANT: JENNINGS, Michael Paul
; APPLICANT: MOXON, E. Richard
; TITLE OF INVENTION: NOVEL SURFACE ANTIGEN
; FILE REFERENCE: 065064/0128
; CURRENT APPLICATION NUMBER: US/09/669,974
; CURRENT FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US 09/377,155
; PRIOR FILING DATE: 1999-08-19
; PRIOR APPLICATION NUMBER: PCT/AU98/01031
; PRIOR FILING DATE: 1998-12-14
; PRIOR APPLICATION NUMBER: GB 9726398.2
; PRIOR FILING DATE: 1997-12-12
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
; LENGTH: 599
; TYPE: PRT
; ORGANISM: Neisseria meningitidis
US-09-669-974-15

Query Match 96.2%; Score 2260; DB 4; Length 599;
Best Local Similarity 96.5%; Pred. No. 3.8e-172;
Matches 445; Conservative 5; Mismatches 11; Indels 0; Gaps 0;

```
Qy          2 DETGLINVETEKLSFGANGKKVNIISDTKGLNFAKETAGTNGDTTVHLNGIGSTLTDMLL 61
              | | | :||||||||||||| |||| ||||||||||||||||||||||||||||||| ||
Db        139 DLTDLTSVETEKLSFGANGNKVNITSDTKGLNFAKETAGTNGDTTVHLNGIGSTLTDTLL 198

Qy         62 NTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLS 121
              ||||||||||||||||:|||||||||||||||||||||||||||||| |||||| ||||||||
Db        199 NTGATTNVTNDNVTDDKKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVHTYDTVEFLS 258

Qy        122 ADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKGKGENGSSTDEGEGLVTAK 181
              |||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||
Db        259 ADTKTTTVNVESKDNGKRTEVKIGAKTSVIKEKDGKLVTGKGKGENGSSTDEGEGLVTAK 318

Qy        182 EVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTKVTFASGNGTTATVSKDDQGNITVK 241
              |||||||||||||||||||||||||||||||| |||||| ||||||||||||||||||||
Db        319 EVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVK 378

Qy        242 YDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEIT 301
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        379 YDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEIT 438

Qy        302 RNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGALNVGSKDANKPVRITNVAPGVKE 361
              |||||||||||||||||||||||||||||||||:||||||||||||||||||||||||||
Db        439 RNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDKGALNVGSKDANKPVRITNVAPGVKE 498

Qy        362 GDVTNVAQLKGVAQNLNNRIDNVNGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYLG 421
```

```
                ||||||||||||||||||||||:||||||||||||||||||||||||||||||||||| |
Db        499 GDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRG 558

Qy        422 EAGYAIGYSSISAGGNWIIKGTASGNSRGHFGASASVGYQW 462
              |||||||||||| |||||||||||||||||||||||||||
Db        559 EAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 599


RESULT 9
US-09-377-155-7
; Sequence 7, Application US/09377155
; Patent No. 6197312
; GENERAL INFORMATION:
;  APPLICANT:  PEAK, Ian Richard Anselm
;  APPLICANT:  JENNINGS, Michael Paul
;  APPLICANT:  MOXON, E. Richard
;  TITLE OF INVENTION:  NOVEL SURFACE ANTIGEN
;  FILE REFERENCE:  065064/0128
;  CURRENT APPLICATION NUMBER: US/09/377,155
;  CURRENT FILING DATE:  1999-08-19
;  PRIOR APPLICATION NUMBER:  PCT/AU98/01031
;  PRIOR FILING DATE:  1998-12-14
;  PRIOR APPLICATION NUMBER:  GB 9726398.2
;  PRIOR FILING DATE:  1997-12-12
;  NUMBER OF SEQ ID NOS:  33
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 7
;   LENGTH: 594
;   TYPE: PRT
;   ORGANISM: Neisseria meningitidis
US-09-377-155-7

  Query Match             95.6%;  Score 2247;  DB 3;  Length 594;
  Best Local Similarity   96.3%;  Pred. No. 4.1e-171;
  Matches  444;  Conservative    2;  Mismatches   15;  Indels    0;  Gaps    0;

Qy          2 DETGLINVETEKLSFGANGKKVNIISDTKGLNFAKETAGTNGDTTVHLNGIGSTLTDMLL 61
              | | | :||||||||||||| |||| |||||||||||||||| |||||||||||| ||
Db        134 DLTDLTSVETEKLSFGANGNKVNITSDTKGLNFAKETAGTNGDPTVHLNGIGSTLTDTLL 193

Qy         62 NTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLS 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        194 NTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLS 253

Qy        122 ADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKGKGENGSSTDEGEGLVTAK 181
              |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        254 ADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKGKDENGSSTDEGEGLVTAK 313

Qy        182 EVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTKVTFASGNGTTATVSKDDQGNITVK 241
              |||||||||||||||||||||||||||||||| |||||| ||||||||||||||||
Db        314 EVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVK 373

Qy        242 YDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEIT 301
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        374 YDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEIT 433

Qy        302 RNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGALNVGSKDANKPVRITNVAPGVKE 361
              |||||||||||| |||||||||||||||||||||||||||| ||||||||||||||||
Db        434 RNGKNIDIATSMAPQFSSVSLGAGADAPTLSVDDEGALNVGSKDTNKPVRITNVAPGVKE 493

Qy        362 GDVTNVAQLKGVAQNLNNRIDNVNGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYLG 421
              |||||||||||||||||||||:||||||||||||||||||||||||||||| || |
Db        494 GDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGDTYRG 553

Qy        422 EAGYAIGYSSISAGGNWIIKGTASGNSRGHFGASASVGYQW 462
              |||||||||||| |||||||||||||||||||||||||||
Db        554 EAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 594


RESULT 10
US-09-669-974-7
; Sequence 7, Application US/09669974
; Patent No. 6333173
; GENERAL INFORMATION:
;  APPLICANT: PEAK, Ian Richard Anselm
;  APPLICANT:  JENNINGS, Michael Paul
;  APPLICANT:  MOXON, E. Richard
;  TITLE OF INVENTION: NOVEL SURFACE ANTIGEN
;  FILE REFERENCE: 065064/0128
;  CURRENT APPLICATION NUMBER: US/09/669,974
;  CURRENT FILING DATE:  2000-09-26
;  PRIOR APPLICATION NUMBER: US 09/377,155
;  PRIOR FILING DATE: 1999-08-19
;  PRIOR APPLICATION NUMBER: PCT/AU98/01031
;  PRIOR FILING DATE: 1998-12-14
;  PRIOR APPLICATION NUMBER: GB 9726398.2
;  PRIOR FILING DATE: 1997-12-12
;  NUMBER OF SEQ ID NOS: 33
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
;   LENGTH: 594
;   TYPE: PRT
;   ORGANISM: Neisseria meningitidis
US-09-669-974-7

  Query Match             95.6%;  Score 2247;  DB 4;  Length 594;
  Best Local Similarity   96.3%;  Pred. No. 4.1e-171;
  Matches  444;  Conservative    2;  Mismatches   15;  Indels    0;  Gaps    0;

Qy          2 DETGLINVETEKLSFGANGKKVNIISDTKGLNFAKETAGTNGDTTVHLNGIGSTLTDMLL 61
              | | | :||||||||||||| |||| |||||||||||||||| |||||||||||| ||
Db        134 DLTDLTSVETEKLSFGANGNKVNITSDTKGLNFAKETAGTNGDPTVHLNGIGSTLTDTLL 193

Qy         62 NTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLS 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        194 NTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLS 253

Qy        122 ADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKGKGENGSSTDEGEGLVTAK 181
              |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        254 ADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKGKDENGSSTDEGEGLVTAK 313

Qy        182 EVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTKVTFASGNGTTATVSKDDQGNITVK 241
              |||||||||||||||||||||||||||||||| |||||| ||||||||||||||||
Db        314 EVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVK 373

Qy        242 YDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEIT 301
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        374 YDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEIT 433

Qy        302 RNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGALNVGSKDANKPVRITNVAPGVKE 361
              |||||||||||| |||||||||||||||||||||||||||| ||||||||||||||||
Db        434 RNGKNIDIATSMAPQFSSVSLGAGADAPTLSVDDEGALNVGSKDTNKPVRITNVAPGVKE 493

Qy        362 GDVTNVAQLKGVAQNLNNRIDNVNGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYLG 421
              |||||||||||||||||||||:||||||||||||||||||||||||||||| || |
Db        494 GDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGDTYRG 553

Qy        422 EAGYAIGYSSISAGGNWIIKGTASGNSRGHFGASASVGYQW 462
              |||||||||||| |||||||||||||||||||||||||||
Db        554 EAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 594


RESULT 11
US-09-377-155-9
; Sequence 9, Application US/09377155
; Patent No. 6197312
; GENERAL INFORMATION:
;  APPLICANT:  PEAK, Ian Richard Anselm
;  APPLICANT:  JENNINGS, Michael Paul
;  APPLICANT:  MOXON, E. Richard
;  TITLE OF INVENTION:  NOVEL SURFACE ANTIGEN
;  FILE REFERENCE:  065064/0128
;  CURRENT APPLICATION NUMBER: US/09/377,155
;  CURRENT FILING DATE:  1999-08-19
;  PRIOR APPLICATION NUMBER:  PCT/AU98/01031
;  PRIOR FILING DATE:  1998-12-14
```

```
;  PRIOR APPLICATION NUMBER:  GB 9726398.2
;  PRIOR FILING DATE:  1997-12-12
;  NUMBER OF SEQ ID NOS:  33
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 9
;   LENGTH: 594
;   TYPE: PRT
;   ORGANISM: Neisseria meningitidis
US-09-377-155-9

  Query Match             94.2%;  Score 2213;  DB 3;  Length 594;
  Best Local Similarity   95.0%;  Pred. No. 2.1e-168;
  Matches  438;  Conservative    5;  Mismatches   18;  Indels    0;  Gaps    0;

Qy          2 DETGLINVETEKLSFGANGKKVNIISDTKGLNFAKETAGTNGDTTVHLNGIGSTLTDMLL 61
                | | | :| |||||| ||  |||| |||||||||:|| |||||||||||||||| ||
Db        134 DLTDLTSVGTEKLSFSANSNKVNITSDTKGLNFAKKTAETNGDTTVHLNGIGSTLTDTLL 193

Qy         62 NTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLS 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        194 NTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLS 253

Qy        122 ADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKGKGENGSSTDEGEGLVTAK 181
              ||||||||||||||||:||||||||||||||||||||||| |||| ||||:|||||||
Db        254 ADTKTTTVNVESKDNGKRTEVKIGAKTSVIKEKDGKLVTGKDKGENDSSTDKGEGLVTAK 313

Qy        182 EVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTKVTFASGNGTTATVSKDDQGNITVK 241
              |||||||||||||||||||||||||||||||||||| |||||| |||||||||||||||
Db        314 EVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVM 373

Qy        242 YDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEIT 301
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        374 YDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEIT 433

Qy        302 RNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGALNVGSKDANKPVRITNVAPGVKE 361
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        434 RNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGALNVGSKDANKPVRITNVAPGVKE 493

Qy        362 GDVTNVAQLKGVAQNLNNRIDNVNGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYLG 421
              ||||||||||||||||| ||||:||||||||||||||||||||||||||||||||||| |
Db        494 GDVTNVAQLKGVAQNLNNHIDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRG 553

Qy        422 EAGYAIGYSSISAGGNWIIKGTASGNSRGHFGASASVGYQW 462
              ||||||||||| |||||||||||||||||||||||||||||
Db        554 EAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 594

RESULT 12
US-09-669-974-9
; Sequence 9, Application US/09669974
; Patent No. 6333173
; GENERAL INFORMATION:
;  APPLICANT: PEAK, Ian Richard Anselm
;  APPLICANT:  JENNINGS, Michael Paul
;  APPLICANT:  MOXON, E. Richard
;  TITLE OF INVENTION: NOVEL SURFACE ANTIGEN
;  FILE REFERENCE: 065064/0128
;  CURRENT APPLICATION NUMBER: US/09/669,974
;  CURRENT FILING DATE:  2000-09-26
;  PRIOR APPLICATION NUMBER: US 09/377,155
;  PRIOR FILING DATE: 1999-08-19
;  PRIOR APPLICATION NUMBER: PCT/AU98/01031
;  PRIOR FILING DATE: 1998-12-14
;  PRIOR APPLICATION NUMBER: GB 9726398.2
;  PRIOR FILING DATE: 1997-12-12
;  NUMBER OF SEQ ID NOS: 33
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
;   LENGTH: 594
;   TYPE: PRT
;   ORGANISM: Neisseria meningitidis
US-09-669-974-9

  Query Match             94.2%;  Score 2213;  DB 4;  Length 594;
  Best Local Similarity   95.0%;  Pred. No. 2.1e-168;
  Matches  438;  Conservative    5;  Mismatches   18;  Indels    0;  Gaps    0;

Qy          2 DETGLINVETEKLSFGANGKKVNIISDTKGLNFAKETAGTNGDTTVHLNGIGSTLTDMLL 61
                | | | :| |||||| ||  |||| |||||||||:|| |||||||||||||||| ||
Db        134 DLTDLTSVGTEKLSFSANSNKVNITSDTKGLNFAKKTAETNGDTTVHLNGIGSTLTDTLL 193

Qy         62 NTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLS 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        194 NTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLS 253

Qy        122 ADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKGKGENGSSTDEGEGLVTAK 181
              ||||||||||||||||:||||||||||||||||||||||| |||| ||||:|||||||
Db        254 ADTKTTTVNVESKDNGKRTEVKIGAKTSVIKEKDGKLVTGKDKGENDSSTDKGEGLVTAK 313

Qy        182 EVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTKVTFASGNGTTATVSKDDQGNITVK 241
              |||||||||||||||||||||||||||||||||||| |||||| |||||||||||||||
Db        314 EVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVM 373

Qy        242 YDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEIT 301
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        374 YDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEIT 433

Qy        302 RNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGALNVGSKDANKPVRITNVAPGVKE 361
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        434 RNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGALNVGSKDANKPVRITNVAPGVKE 493

Qy        362 GDVTNVAQLKGVAQNLNNRIDNVNGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYLG 421
              ||||||||||||||||| ||||:||||||||||||||||||||||||||||||||||| |
Db        494 GDVTNVAQLKGVAQNLNNHIDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRG 553

Qy        422 EAGYAIGYSSISAGGNWIIKGTASGNSRGHFGASASVGYQW 462
              ||||||||||| |||||||||||||||||||||||||||||
Db        554 EAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 594

RESULT 13
US-09-377-155-11
; Sequence 11, Application US/09377155
; Patent No. 6197312
; GENERAL INFORMATION:
;  APPLICANT:  PEAK, Ian Richard Anselm
;  APPLICANT:  JENNINGS, Michael Paul
;  APPLICANT:  MOXON, E. Richard
;  TITLE OF INVENTION:  NOVEL SURFACE ANTIGEN
;  FILE REFERENCE:  065064/0128
;  CURRENT APPLICATION NUMBER: US/09/377,155
;  CURRENT FILING DATE:  1999-08-19
;  PRIOR APPLICATION NUMBER:  PCT/AU98/01031
;  PRIOR FILING DATE:  1998-12-14
;  PRIOR APPLICATION NUMBER:  GB 9726398.2
;  PRIOR FILING DATE:  1997-12-12
;  NUMBER OF SEQ ID NOS:  33
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 11
;   LENGTH: 591
;   TYPE: PRT
;   ORGANISM: Neisseria meningitidis
US-09-377-155-11

  Query Match             94.1%;  Score 2210.5;  DB 3;  Length 591;
  Best Local Similarity   95.4%;  Pred. No. 3.3e-168;
  Matches  440;  Conservative    3;  Mismatches   17;  Indels    1;  Gaps    1;

Qy          2 DETGLINVETEKLSFGANGKKVNIISDTKGLNFAKETAGTNGDTTVHLNGIGSTLTDMLL 61
                | | | :| |||||| ||| |||| |||||||||||||||||||||||||||||| ||
Db        132 DLTDLTSVGTEKLSFSANGNKVNITSDTKGLNFAKETAGTNGDTTVHLNGIGSTLTDTLL 191

Qy         62 NTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLS 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        192 NTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLS 251
```

```
Qy        122 ADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKGKGENGSSTDEGEGLVTAK 181
              ||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        252 ADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAK 311

Qy        182 EVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTKVTFASGNGTTATVSKDDQGNITVK 241
              |||||||||||||||||||||||||||||||| |||||| ||||||||||||||||||
Db        312 EVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVM 371

Qy        242 YDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEIT 301
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        372 YDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEIT 431

Qy        302 RNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGALNVGSKDANKPVRITNVAPGVKE 361
              |||||||||||||||||||||||||||||||| : ||||||| ||||||||||||||||
Db        432 RNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGD-ALNVGSKKDNKPVRITNVAPGVKE 490

Qy        362 GDVTNVAQLKGVAQNLNNRIDNVNGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYLG 421
              |||||||||||||||||||||:||||||||||||||||||||||||||||||||||| |
Db        491 GDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRG 550

Qy        422 EAGYAIGYSSISAGGNWIIKGTASGNSRGHFGASASVGYQW 462
              |||||||||||| ||||||||||||||||||||||||||||
Db        551 EAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 591


RESULT 14
US-09-377-155-21
; Sequence 21, Application US/09377155
; Patent No. 6197312
; GENERAL INFORMATION:
;  APPLICANT:  PEAK, Ian Richard Anselm
;  APPLICANT:  JENNINGS, Michael Paul
;  APPLICANT:  MOXON, E. Richard
;  TITLE OF INVENTION:  NOVEL SURFACE ANTIGEN
;  FILE REFERENCE:  065064/0128
;  CURRENT APPLICATION NUMBER: US/09/377,155
;  CURRENT FILING DATE:  1999-08-19
;  PRIOR APPLICATION NUMBER:  PCT/AU98/01031
;  PRIOR FILING DATE:  1998-12-14
;  PRIOR APPLICATION NUMBER:  GB 9726398.2
;  PRIOR FILING DATE:  1997-12-12
;  NUMBER OF SEQ ID NOS:  33
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 21
;   LENGTH: 591
;   TYPE: PRT
;   ORGANISM: Neisseria meningitidis
US-09-377-155-21

  Query Match             94.1%;  Score 2210.5;  DB 3;  Length 591;
  Best Local Similarity   95.4%;  Pred. No. 3.3e-168;
  Matches  440;  Conservative    3;  Mismatches   17;  Indels    1;  Gaps    1;

Qy          2 DETGLINVETEKLSFGANGKKVNIISDTKGLNFAKETAGTNGDTTVHLNGIGSTLTDMLL 61
              | | | :| |||||| ||| |||| ||||||||||||||||||||||||||||||||| ||
Db        132 DLTDLTSVGTEKLSFSANGNKVNITSDTKGLNFAKETAGTNGDTTVHLNGIGSTLTDTLL 191

Qy         62 NTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLS 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        192 NTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLS 251

Qy        122 ADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKGKGENGSSTDEGEGLVTAK 181
              |||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        252 ADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAK 311

Qy        182 EVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTKVTFASGNGTTATVSKDDQGNITVK 241
              |||||||||||||||||||||||||||||||| |||||| ||||||||||||||||||
Db        312 EVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVM 371

Qy        242 YDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEIT 301
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        372 YDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEIT 431

Qy        302 RNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGALNVGSKDANKPVRITNVAPGVKE 361
              |||||||||||||||||||||||||||||||| : ||||||| ||||||||||||||||
Db        432 RNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGD-ALNVGSKKDNKPVRITNVAPGVKE 490

Qy        362 GDVTNVAQLKGVAQNLNNRIDNVNGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYLG 421
              |||||||||||||||||||||:||||||||||||||||||||||||||||||||||| |
Db        491 GDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRG 550

Qy        422 EAGYAIGYSSISAGGNWIIKGTASGNSRGHFGASASVGYQW 462
              |||||||||||| ||||||||||||||||||||||||||||
Db        551 EAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 591


RESULT 15
US-09-669-974-11
; Sequence 11, Application US/09669974
; Patent No. 6333173
; GENERAL INFORMATION:
;  APPLICANT: PEAK, Ian Richard Anselm
;  APPLICANT:  JENNINGS, Michael Paul
;  APPLICANT:  MOXON, E. Richard
;  TITLE OF INVENTION: NOVEL SURFACE ANTIGEN
;  FILE REFERENCE: 065064/0128
;  CURRENT APPLICATION NUMBER: US/09/669,974
;  CURRENT FILING DATE:  2000-09-26
;  PRIOR APPLICATION NUMBER: US 09/377,155
;  PRIOR FILING DATE: 1999-08-19
;  PRIOR APPLICATION NUMBER: PCT/AU98/01031
;  PRIOR FILING DATE: 1998-12-14
;  PRIOR APPLICATION NUMBER: GB 9726398.2
;  PRIOR FILING DATE: 1997-12-12
;  NUMBER OF SEQ ID NOS: 33
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
;   LENGTH: 591
;   TYPE: PRT
;   ORGANISM: Neisseria meningitidis
US-09-669-974-11

  Query Match             94.1%;  Score 2210.5;  DB 4;  Length 591;
  Best Local Similarity   95.4%;  Pred. No. 3.3e-168;
  Matches  440;  Conservative    3;  Mismatches   17;  Indels    1;  Gaps    1;

Qy          2 DETGLINVETEKLSFGANGKKVNIISDTKGLNFAKETAGTNGDTTVHLNGIGSTLTDMLL 61
              | | | :| |||||| ||| |||| ||||||||||||||||||||||||||||||||| ||
Db        132 DLTDLTSVGTEKLSFSANGNKVNITSDTKGLNFAKETAGTNGDTTVHLNGIGSTLTDTLL 191

Qy         62 NTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLS 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        192 NTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLS 251

Qy        122 ADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKGKGENGSSTDEGEGLVTAK 181
              |||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        252 ADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAK 311

Qy        182 EVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTKVTFASGNGTTATVSKDDQGNITVK 241
              |||||||||||||||||||||||||||||||| |||||| ||||||||||||||||||
Db        312 EVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVM 371

Qy        242 YDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEIT 301
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        372 YDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEIT 431

Qy        302 RNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGALNVGSKDANKPVRITNVAPGVKE 361
              |||||||||||||||||||||||||||||||| : ||||||| ||||||||||||||||
Db        432 RNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGD-ALNVGSKKDNKPVRITNVAPGVKE 490

Qy        362 GDVTNVAQLKGVAQNLNNRIDNVNGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYLG 421
              |||||||||||||||||||||:||||||||||||||||||||||||||||||||||| |
Db        491 GDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRG 550

Qy        422 EAGYAIGYSSISAGGNWIIKGTASGNSRGHFGASASVGYQW 462
              |||||||||||| ||||||||||||||||||||||||||||
Db        551 EAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 591
```

Db 551 EAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 591

Search completed: October 6, 2003, 09:35:58
Job time : 14.3548 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: October 6, 2003, 09:13:35 ; Search time 13.2249 Seconds
(without alignments)
3359.577 Million cell updates/sec

Title: US-09-771-382-36
Perfect score: 2350
Sequence: 1 TDETGLINVETEKLSFGANG.....TASGNSRGHFGASASVGYQW 462

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_76:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2210.5	94.1	591	2	G81133	adhesin NMB0992 [i
2	2136	90.9	592	2	A81888	probable surface f
3	390.5	16.6	298	2	I64138	adhesin homolog HI
4	360	15.3	1588	2	A86036	probable adhesin Z
5	360	15.3	1588	2	H91188	probable adhesin E
6	359.5	15.3	2059	2	D82671	surface protein XF
7	357	15.2	1190	2	A82615	surface protein XF
8	348.5	14.8	1107	2	AC0976	probable autotrans
9	317	13.5	658	2	AH0110	probable surface p
10	212.5	9.0	1004	2	C82672	surface-exposed ou
11	207	8.8	3705	2	AD0123	probable autotrans
12	200.5	8.5	1910	2	AF0394	probable adhesin h
13	195	8.3	1536	2	A43855	high-molecular-wei
14	190	8.1	4919	2	T31105	hypothetical prote
15	188.5	8.0	936	2	I40711	sapB protein - Cam
16	187.5	8.0	365	2	AB3486	cell surface prote
17	187.5	8.0	2249	2	A41477	190K surface antig
18	187	8.0	2020	2	C48399	ABC-type transport
19	186	7.9	1651	2	JC1340	outer membrane pro
20	184.5	7.9	5188	2	B85547	probable RTX famil
21	183.5	7.8	5291	2	F90696	hypothetical prote
22	181	7.7	1109	2	A56143	surface-array prot
23	181	7.7	2551	2	B98047	hypothetical prote
24	180.5	7.7	1477	2	B43855	high-molecular-wei
25	180.5	7.7	1655	2	E97835	hypothetical prote
26	180	7.7	1361	2	T03415	S-layer protein -
27	179.5	7.6	364	2	A81019	adhesin/invasin, p
28	178.5	7.6	3029	2	S76109	hypothetical prote
29	178	7.6	4152	2	T31102	filamentous hemagg
30	177	7.5	1091	2	G64964	hypothetical prote
31	177	7.5	2468	2	A83412	hypothetical prote
32	176.5	7.5	893	2	A37284	surface-array prot
33	175.5	7.5	1275	2	T33369	hypothetical prote
34	174	7.4	585	2	F85809	hypothetical prote
35	173.5	7.4	1577	2	A35140	hemolysin A precur
36	173	7.4	1567	2	S11672	ice nucleation pro
37	172.5	7.3	1035	2	AD3203	autotransporter pr
38	172	7.3	585	2	F90961	flagellin [importe
39	171.5	7.3	1268	2	B99789	hemagglutinin/hemo
40	171.5	7.3	1270	2	E85649	hypothetical prote
41	171.5	7.3	1300	2	S07575	outer membrane pro
42	171.5	7.3	1430	2	AF0351	probable autotrans
43	171	7.3	978	2	D81411	probable lipoprote
44	171	7.3	3013	2	AB0480	probable invasin Y
45	170	7.2	1052	2	AF2959	conserved hypothet

ALIGNMENTS

RESULT 1
G81133
adhesin NMB0992 [imported] - Neisseria meningitidis (strain MC58 serogroup B)
C;Species: Neisseria meningitidis
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001
C;Accession: G81133
R;Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B. ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Masignani, V.; Pizza, M.
Science 287, 1809-1815, 2000
A;Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.;
A;Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A;Reference number: A81000; MUID:20175755; PMID:10710307
A;Accession: G81133
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-591 <TET>
A;Cross-references: GB:AE002450; GB:AE002098; NID:g7226229; PIDN:AAF41395.1; PID:g722
A;Experimental source: serogroup B, strain MC58
C;Genetics:
A;Gene: NMB0992

Query Match 94.1%; Score 2210.5; DB 2; Length 591;
Best Local Similarity 95.4%; Pred. No. 9.3e-110;
Matches 440; Conservative 3; Mismatches 17; Indels 1; Gaps 1;

```
Qy          2 DETGLINVETEKLSFGANGKKVNIISDTKGLNFAKETAGTNGDTTVHLNGIGSTLTDMLL 61
              | | | :| |||||| ||| |||| ||||||||||||||||||||||||||||||| ||
Db        132 DLTDLTSVGTEKLSFSANGNKVNITSDTKGLNFAKETAGTNGDTTVHLNGIGSTLTDTLL 191

Qy         62 NTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLS 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        192 NTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLS 251

Qy        122 ADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKGKGENGSSTDEGEGLVTAK 181
              ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        252 ADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAK 311

Qy        182 EVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTKVTFASGNGTTATVSKDDQGNITVK 241
              |||||||||||||||||||||||||||||||| |||||| ||||||||||||||||
Db        312 EVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVM 371

Qy        242 YDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEIT 301
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        372 YDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEIT 431

Qy        302 RNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGALNVGSKDANKPVRITNVAPGVKE 361
              |||||||||||||||||||||||||||||||| : ||||||| ||||||||||||||||
Db        432 RNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGD-ALNVGSKKDNKPVRITNVAPGVKE 490

Qy        362 GDVTNVAQLKGVAQNLNNRIDNVNGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYLG 421
              |||||||||||||||||||:|||||||||||||||-|||||||||||||||||||||| |
```

```
Db        491 GDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRG 550

Qy        422 EAGYAIGYSSISAGGNWIIKGTASGNSRGHFGASASVGYQW 462
              ||||||||||| |||||||||||||||||||||||||||||
Db        551 EAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 591


RESULT 2
A81888
probable surface fibril protein NMA1200 [imported] - Neisseria meningitidis (strain Z249
C;Species: Neisseria meningitidis
C;Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001
C;Accession: A81888
R;Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel
; Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,
Nature 404, 502-506, 2000
A;Title: Complete DNA sequence of a serogroup A strain of Neisseria menigitidis Z2491.
A;Reference number: A81775; MUID:20222556; PMID:10761919
A;Accession: A81888
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-592 <PAR>
A;Cross-references: GB:AL162755; GB:AL157959; NID:g7379742; PIDN:CAB84461.1; PID:g737989
A;Experimental source: serogroup A, strain Z2491
C;Genetics:
A;Gene: NMA1200

  Query Match             90.9%;  Score 2136;  DB 2;  Length 592;
  Best Local Similarity   92.2%;  Pred. No. 8.1e-106;
  Matches  427;  Conservative    9;  Mismatches   23;  Indels    4;  Gaps    2;

Qy          2 DETGLINVETEKLSFGANGKKVNIISDTKGLNFAKETAGTNGDTTVHLNGIGSTLTDMLL 61
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db        132 DLTGLINVETEKLSFGANGKKVNIISDTKGLNFAKETAGTNGDTTVHLNGIGSTLTDTLA 191

Qy         62 NTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPGTTA--SDNVDFVRTYDTVEF 119
                : |:     :  |     ||||:|||||||||||||| |:|   |:|||||||||||||
Db        192 GSSASHVDAGNQST--HYTRAASIKDVLNAGWNIKGVKTGSTTGQSENVDFVRTYDTVEF 249

Qy        120 LSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKGKGENGSSTDEGEGLVT 179
              ||||||||||||||||||:|||||||||||||||||||||||||||||||||||||||||
Db        250 LSADTKTTTVNVESKDNGKRTEVKIGAKTSVIKEKDGKLVTGKGKGENGSSTDEGEGLVT 309

Qy        180 AKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTKVTFASGNGTTATVSKDDQGNIT 239
              |||||||||||||||||||||||||||||||||||||| |||||| ||||||||||||||
Db        310 AKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNIT 369

Qy        240 VKYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIE 299
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        370 VMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIE 429

Qy        300 ITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGALNVGSKDANKPVRITNVAPGV 359
              |:|||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        430 ISRNGKNIDIATSMAPQFSSVSLGAGADAPTLSVDDEGALNVGSKDANKPVRITNVAPGV 489

Qy        360 KEGDVTNVAQLKGVAQNLNNRIDNVNGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTY 419
              |||||||||||||||||||||||:|||||||||||||||||||||||||||||||||||||
Db        490 KEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTY 549

Qy        420 LGEAGYAIGYSSISAGGNWIIKGTASGNSRGHFGASASVGYQW 462
              |||||||||||| |||||||||||||||||||||||||||||
Db        550 RGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 592


RESULT 3
I64138
adhesin homolog HI1732 - Haemophilus influenzae (strain Rd KW20)
C;Species: Haemophilus influenzae
C;Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 24-Oct-1997
C;Accession: I64138
R;Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A
; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J
, D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M.
Science 269, 496-512, 1995
A;Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Vente
A;Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
A;Reference number: A64000; MUID:95350630; PMID:7542800
A;Accession: I64138
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-298 <TIGR>
A;Cross-references: GB:U32846; GB:L42023; NID:g1574588; PID:g1574589; TIGR:HI1732

  Query Match             16.6%;  Score 390.5;  DB 2;  Length 298;
  Best Local Similarity   48.9%;  Pred. No. 6.9e-14;
  Matches   91;  Conservative   24;  Mismatches   48;  Indels   23;  Gaps    5;

Qy          7 INVETEKLS----------FGANGKKVNIISDTKGLNFAKETAGTNGDTTVHLNGIGST 55
              ::|:| |:|           |    ||:| |   ||  ||   |||||| ||||| ||
Db        114 LDVKTAKVSDTLTIGGNTPAAGGATPKVSITSTADGLKLAK---GTNGDTAVHLNGLAST 170

Qy         56 LTDMLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYD 115
              | |:  ||||:|:||  : :| || |||::||||||||||| |     ::||| |  ||
Db        171 LPDVTTNTGASTSVT-FSPSDIEKTRAATIKDVLNAGWNIKGAKVAGGNTENVDLVAGYD 229

Qy        116 TVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGK-------GKGENG 168
               |||:: |  |  | : :|:||| ||||   ||||||: :|||:|||       |   |
Db        230 NVEFITGDKNTLDVVLTAKENGKTTEVKFTPKTSVIKDNNGKLLTGKQLKDANTGTATNA 289

Qy        169 S-STDE 173
              :  |||
Db        290 TEDTDE 295


RESULT 4
A86036
probable adhesin Z5029 [imported] - Escherichia coli (strain O157:H7, substrain EDL93
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C;Accession: A86036
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apoda
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: A86036
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1588 <STO>
A;Cross-references: GB:AE005174; NID:g12518349; PIDN:AAG58749.1; GSPDB:GN00145; UWGP:
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: Z5029

  Query Match             15.3%;  Score 360;  DB 2;  Length 1588;
  Best Local Similarity   24.9%;  Pred. No. 2.2e-11;
  Matches  141;  Conservative   73;  Mismatches  236;  Indels  116;  Gaps   18;

Qy          4 TGLINV---ETEKLSFGANGKKVNII-SDTKGLNFAKETAGTNGDTTVHLNGI------- 52
              | :||     |:      ||  :| : ::  || |   :|   | | :  | :
Db       1032 TQIINQLAGNTDATYIQENGAGINYVRTNDDGLAFNDASAQGVGATAIGYNSVAKGDSSV 1091

Qy         53 -----------------GSTLTDMLLNTGA-TTNVTNDNV------TDDEKKRAASVKD- 87
                              |:::  ::  |:  |::| : |      || |   | |: |
Db       1092 AIGQGSYSDVDTGIALGSSSVSSRVIAKGSRDTSITENGVVIGYDTTDGELLGALSIGDD 1151

Qy         88 -----VLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGK---- 138
                  ::|       |  |: | | |  ||          | | |   :  | :
Db       1152 GKYRQIIN-------VADGSEAHDAVT-VRQLQNAIGAVATTPTKYFHANSTEEDSLAVG 1203

Qy        139 KTEVKIGAKTSVIKEK----------DGKLVTGKGKGE------------NGSSTDEGE 175
                : :|||| |  :|          |   : |   |              |||:|  |
Db       1204 TDSLAMGAKTIVNGDKGIGIGYGAYVDANALNGIAIGSNAQVIHVNSIAIGNGSTTTRGA 1263

Qy        176 GLVTAKEVIDAVNKAGWRMKTTTANGQ-------TGQADKFETVTSGTKVTFASGNGTTA 228
                      :||   :       :|:||        | ||         ||| |  :  |
```

```
Db       1264 QTNYTAYNMDAPQNSVGEFSVGSADGQRQITNVAAGSADTDAVNVGQLKVTDAQVSQNTQ 1323

Qy        229 TVSKDDQ-----GNITVKYDVNVGDAL----------NVNQLQNSGWNLDSKAVAGSSGK 273
              :::  |       :     :  :|| :          | : :  |    || |:   |
Db       1324 SITNLDNRVTNLDSRVTNIENGIGDIVTTGSTKYFKTNTDGVDASAQGKDSVAIGSGSIA 1383

Qy        274 VISGNVSPSKGKM---DETVNINAGNNIEITRN---GKNIDIATSMTPQFSSVSLGAGAD 327
                :|:   | :   : |::: :  |     |   |||   | ::    || : |   |
Db       1384 AADNSVALGTGSVATEENTISVGSSTNQRRITNVAAGKNATDAVNVAQLKSSEAGGVRYD 1443

Qy        328 APTLSVDDEGALNVGSKDANKPVRITNVAPGVKEGDVTNVAQLKGVAQ----------N 376
                    |   : :|  :     ||:||: ||   || | ||||   |
Db       1444 TKADGSIDYSNITLGGGNGG-TTRISNVSAGVNNNDVVNYAQLKQSVQETKQYTDQRMVE 1502

Qy        377 LNNRIDNVNGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYLGEAGYAIGYSSISAGG 436
              ::|::          ||| |:|  || ||| || || :|||||| ||:  |:| | :|| |
Db       1503 MDNKLSKTESKLSGGIASAMAMTGLPQAYTPGASMASIGGGTYNGESAVALGVSMVSANG 1562

Qy        437 NWIIKGTASGNSRGHFGASASVGYQW 462
               |: |   | ||:| : |:   | ||
Db       1563 RWVYKLQGSTNSQGEYSAALGAGIQW 1588


RESULT 5
H91188
probable adhesin ECs4480 [similarity] - Escherichia coli (strain O157:H7, substrain RIMD
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 02-Nov-2001
C;Accession: H91188
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genc
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: H91188
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1588 <HAY>
A;Cross-references: GB:BA000007; PIDN:BAB37903.1; PID:g13363955; GSPDB:GN00154
A;Experimental source: strain O157:H7, substrain RIMD 0509952
C;Genetics:
A;Gene: ECs4480

  Query Match             15.3%;  Score 360;  DB 2;  Length 1588;
  Best Local Similarity   24.9%;  Pred. No. 2.2e-11;
  Matches  141;  Conservative   73;  Mismatches  236;  Indels  116;  Gaps   18;

Qy          4 TGLINV---ETEKLSFGANGKKVNII-SDTKGLNFAKETAGTNGDTTVHLNGI------- 52
              | :||     |:      ||  :| : ::  || |   :|   | | :  | :
Db       1032 TQIINQLAGNTDATYIQENGAGINYVRTNDDGLAFNDASAQGVGATAIGYNSVAKGDSSV 1091

Qy         53 -----------------GSTLTDMLLNTGA-TTNVTNDNV------TDDEKKRAASVKD- 87
                              |:::  ::  |:  |::| : |       || |   | |: |
Db       1092 AIGQGSYSDVDTGIALGSSSVSSRVIAKGSRDTSITENGVVIGYDTTDGELLGALSIGDD 1151

Qy         88 -----VLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGK---- 138
                  ::|       |  |: | | |  ||         | | |   :  | :
Db       1152 GKYRQIIN-------VADGSEAHDAVT-VRQLQNAIGAVATTPTKYFHANSTEEDSLAVG 1203

Qy        139 KTEVKIGAKTSVIKEK----------DGKLVTGKGKGE-------------NGSSTDEGE 175
               : :|||| |  :|          |   : |   |                |||:|  |
Db       1204 TDSLAMGAKTIVNGDKGIGIGYGAYVDANALNGIAIGSNAQVIHVNSIAIGNGSTTTRGA 1263

Qy        176 GLVTAKEVIDAVNKAGWRMKTTTANGQ-------TGQADKFETVTSGTKVTFASGNGTTA 228
                    :||   :        :|:||         | ||          ||| |  :  |
Db       1264 QTNYTAYNMDAPQNSVGEFSVGSADGQRQITNVAAGSADTDAVNVGQLKVTDAQVSQNTQ 1323

Qy        229 TVSKDDQ-----GNITVKYDVNVGDAL----------NVNQLQNSGWNLDSKAVAGSSGK 273
              :::  |       :     :  :|| :          | : :  |    || |:   |
Db       1324 SITNLDNRVTNLDSRVTNIENGIGDIVTTGSTKYFKTNTDGVDASAQGKDSVAIGSGSIA 1383

Qy        274 VISGNVSPSKGKM---DETVNINAGNNIEITRN---GKNIDIATSMTPQFSSVSLGAGAD 327
                :|:   | :   : |::: :  |     |   |||   | ::    || : |   |
Db       1384 AADNSVALGTGSVATEENTISVGSSTNQRRITNVAAGKNATDAVNVAQLKSSEAGGVRYD 1443

Qy        328 APTLSVDDEGALNVGSKDANKPVRITNVAPGVKEGDVTNVAQLKGVAQ----------N 376
                    |   : :|  :     ||:||: ||   || | ||||   |
Db       1444 TKADGSIDYSNITLGGGNGG-TTRISNVSAGVNNNDVVNYAQLKQSVQETKQYTDQRMVE 1502

Qy        377 LNNRIDNVNGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYLGEAGYAIGYSSISAGG 436
              ::|::          ||| |:|  || ||| || || :|||||| ||:  |:| | :|| |
Db       1503 MDNKLSKTESKLSGGIASAMAMTGLPQAYTPGASMASIGGGTYNGESAVALGVSMVSANG 1562

Qy        437 NWIIKGTASGNSRGHFGASASVGYQW 462
               |: |   | ||:| : |:   | ||
Db       1563 RWVYKLQGSTNSQGEYSAALGAGIQW 1588


RESULT 6
D82671
surface protein XF1529 [imported] - Xylella fastidiosa (strain 9a5c)
C;Species: Xylella fastidiosa
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C;Accession: D82671
R;anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Seq
Nature 406, 151-157, 2000
A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A;Reference number: A82515; MUID:20365717; PMID:10910347
A;Note: for a complete list of authors see reference number A59328 below
A;Accession: D82671
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-2059 <SIM>
A;Cross-references: GB:AE003982; GB:AE003849; NID:g9106554; PIDN:AAF84338.1; GSPDB:GN
A;Experimental source: strain 9a5c
R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer
as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Fr
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; La
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins
A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.
, F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri,
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawa
A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silv
M.; Tsuhako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.
A;Reference number: A59328
A;Contents: annotation
C;Genetics:
A;Gene: XF1529

  Query Match             15.3%;  Score 359.5;  DB 2;  Length 2059;
  Best Local Similarity   24.5%;  Pred. No. 3.3e-11;
  Matches  151;  Conservative   82;  Mismatches  188;  Indels  195;  Gaps   29;

Qy          8 NVETEKLSFGANGKKV--NIISDTKGLNFAKETAGT------NGDTTVH---LNGIGSTL 56
              | : :|     :| ||  |:: |:  |      ::  :      ||:: |:   :|| |  :
Db       1478 NTQLDK-----DGVKVSSNVLLDSNELVITSHSSTSSVKTLANGESVVNRTVVNGDGVNI 1532

Qy         57 TDML-------------------LNTGA--TTNVT-----NDNVTDDEKKRAASVKDVLN 90
                |::                 :| |:   ||||      | |   : |   || : ::
Db       1533 DDVVVVNDLGLSIVGGASLTLSGINAGSHKITNVTAGTEDTDAVNFSQLK---SVSEAVD 1589

Qy         91 AGWNI--------KGVKPGTTASDNVD-------FVRTYDTVEFLSADTK-------TTT 128
                || :        | |  ||    | |          : | |  || | |         |
Db       1590 KGWTLTASGANGSKVVSGGTVDLKNTDGNLAISKSGDSNDVVFNLSKDFKVDEVTAGNTV 1649

Qy        129 VNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKGKGEN-------GSSTDEGEGLVTAK 181
              || :    |. ::| :||    |         :|   |:        | :  :   :  |
Db       1650 VNTDGVKVG--SDVSLGAMGLFIANGPSVTASGFNAGDKVISHVAVGMADTDAVNVSQLK 1707

Qy        182 EVIDAVNKAGWRMKTTTANGQTG--------QADKFETVTSGTKVT-----------FAS 222
              : : :|     |  :|   |  |           |      ||: :            ||
Db       1708 QAVQSVTVKATRYYSTNDGGTQGGNYDGDGATGSKAIAAGVGTQASGEGAAAVGSGAAAS 1767
```

```
Qy        223 GNGTT-----ATVSKD----------DQG----NITVKYD---------VNVG------- 247
              | |:|      |  | |          | |    : | ||         |:||
Db       1768 GKGSTAIGRNAIASADGSVALGDGAKDGGRGAESYTGKYSGVQNNTVGTVSVGDAAKGET 1827

Qy        248 ------------DALNVNQL----QNSGWNLDS-----------KAVAGSSGKVISGN 278
                        ||:|: ||    | |    |              |   | |   : |
Db       1828 RSISNVADAKEAMDAVNLRQLDAVAQKSNLQTDDMRHEINNIEDVFKITKGDSASSVKG- 1886

Qy        279 VSPSKGKMDETVNINA---GNNIEITRN-----GKNIDIATSMTPQFSSVSLGAGADAPT 330
                      : :||   | |  ::       ||| :::       ::|::| |: |
Db       1887 -----------MGVNAMAIGTNAAVSGTESVALGKNTNVSAD-----NAVAIGNGSVA-- 1928

Qy        331 LSVDDEGALNVGSKDANKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVNGNAR- 389
                |   :::|||  : :  ::|||| |  : |  ||:||           | : || |
Db       1929 ---DRANSVSVGSGGSER--QVTNVAAGTADTDAVNVSQLNQGLITAKQYTDGMVGNLRR 1983

Qy        390 ---AGIAQAIATAGLVQAYLPGKSMMAIGGGTYLGEAGYAIGYSSISAGGNWIIKGTASG 446
                 |:| ||||| | |||: |: | ::|  :| |::  |:| |::|  |:|: | : |
Db       1984 ETSGGVAAAIATANLPQAYVQGRGMTSVGVSSYQGQSAIAVGVSAVSESGHWVFKFSGSA 2043

Qy        447 NSRGHFGASASVGYQW 462
                |:| | |  | |||||
Db       2044 NTRSHVGVGAGVGYQW 2059
```

RESULT 7
A82615
surface protein XF1981 [imported] - Xylella fastidiosa (strain 9a5c)
C;Species: Xylella fastidiosa
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C;Accession: A82615
R;anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000
A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A;Reference number: A82515; MUID:20365717; PMID:10910347
A;Note: for a complete list of authors see reference number A59328 below
A;Accession: A82615
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1190 <SIM>
A;Cross-references: GB:AE004017; GB:AE003849; NID:g9107083; PIDN:AAF84783.1; GSPDB:GN001
A;Experimental source: strain 9a5c
R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H
as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigr
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E
A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
, F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
M.; Tsuhako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A;Reference number: A59328
A;Contents: annotation
C;Genetics:
A;Gene: XF1981

  Query Match   15.2%; Score 357; DB 2; Length 1190;
  Best Local Similarity   25.3%; Pred. No. 2.3e-11;
  Matches 132; Conservative   72; Mismatches 189; Indels 128; Gaps   19;

```
Qy         14 LSFGAN------GKKVNIISDTKGLNFAKETAG---------------TNGDTTVHLNG 51
              |: |||      |: |:: :    :  :||:                 | ||| :  ||
Db        726 LASGANSSNVAPGESVDLKNTDGNIVISKESGSNDVLFNLSSSLKLDKLTVGDTVMTTNG 785

Qy         52 IGSTLTDMLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAG-WNIKGVKPGTTASD---- 106
              :       : :| |       :||     :: :    |||   |  |  ||  :|
Db        786 V-------TVGSGVTLGSMGLVITDGPSVTSSGI----NAGSQKITNVAAGTADTDAVNL 834

Qy        107 ------------NVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKE 153
                         :| :  |||        |:    | :     :   | :|   |
```

```
Db        835 SQLNTAMAGSGAKSVHYYSTYD------GGTQGGNYNGDGATGTRSIAVGVGTLASA--- 885

Qy        154 KDGKLVTGKGKGENG-SSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETV 212
               :|    | |   :|  ||  |   | : :   |:   | :     |   ||:    :
Db        886 -EGATAVGSGAAASGKGSTAIGRNAVASADGSVALGD-GAKDGARGAESYTGKYSGLQNN 943

Qy        213 TSGTKVTFASGNGTTATVSKDDQGNITVKYDVNVGDALNVNQL----QNSGWNLDSKAVA 268
              | ||      :  | | |||     |:    :    ||:|: ||    |::   :|:|  :
Db        944 TVGTVSVGDASKGETRTVS-----NVADAKEAT--DAVNLRQLDRVAQDANRYVDNKIES 996

Qy        269 GSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADA 328
               | |:                 | :|: ||            || :     : ::| || |
Db        997 LSEGQTF-------------VKVNSLNN-----------SATPIAAGVDATAIGVGATA 1031

Qy        329 -----------PTLSVDDEGAL------------NVGSKDANKPVRITNVAPGVKEGDVT 365
                        : | |:  |:            :|||  : :  ::|||| |  : |
Db       1032 SGADSIAMGNKASASADNAVAIGNHSVADRANTVSVGSAGSER--QVTNVAAGTADTDAV 1089

Qy        366 NVAQLKGVAQNLNNRIDNVNGNAR----AGIAQAIATAGLVQAYLPGKSMMAIGGGTYLG 421
              ||:||           | | |: |     |:| ||||| | |||:||: | ::|  :| |
Db       1090 NVSQLNQGLITAKQYTDGVVGSLRRDTDGGVAAAIATANLPQAYIPGRGMTSVGVSSYRG 1149

Qy        422 EAGYAIGYSSISAGGNWIIKGTASGNSRGHFGASASVGYQW 462
              ::  |:| ||:|  | |: | : | |:|   |  | |||||
Db       1150 QSAIAVGVSSVSESGRWVFKFSGSANTRSQVGIGAGVGYQW 1190
```

RESULT 8
AC0976
probable autotransporter sapB [imported] - Salmonella enterica subsp. enterica serova
C;Species: Salmonella enterica subsp. enterica serovar Typhi
A;Note: this species has also been called Salmonella typhi
C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C;Accession: AC0976
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Church
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farr
, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens,
A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica se
A;Reference number: AB0502; MUID:21534947; PMID:11677608
A;Accession: AC0976
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1107 <PAR>
A;Cross-references: GB:AL513382; PIDN:CAD03303.1; PID:g16504923; GSPDB:GN00176
C;Genetics:
A;Gene: sapB

  Query Match   14.8%; Score 348.5; DB 2; Length 1107;
  Best Local Similarity   22.0%; Pred. No. 5.8e-11;
  Matches 153; Conservative   88; Mismatches 211; Indels 243; Gaps   23;

```
Qy          1 TDETGLINVETEKLSFGANGKKVNIISDTKGLNFAKETAGTNG----DTTVHLNGIGSTL 56
              ||   | : :|   |    |    |   : |      |   ||    :|   ::   | :
Db        423 TDNALLWDEDTGAFSANHGGSTSKITNVAAGALSEDSTDAVNGSQLYETNQKVDQNTSAI 482

Qy         57 TDMLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFV---RT 113
               |  :|| : ||:  | :: |::: | |      :    |  |  | |||: | :   :
Db        483 AD--INT-SITNLGTDALSWDDEEGAFSASHGTSGTNKITNVAAGEIASDSTDAINGSQL 539

Qy        114 YDT----------VEFLSADTKTTTVN--------VESKDNGKKTEVKIGAKTSVIKEKD 155
              |:|          :  |: ||  | :         : : ||| : :
Db        540 YETNMLISQYNESISQLAGDTSETYITENGTGVKYIRTNDNGLEGQDAYATGNGATAVGY 599

Qy        156 GKLVTGKG---KGENGSSTDEG------------------------EGLV-----TAKE 182
               : :| |    |:| ||: ||                        :|:|     | :|
Db        600 DAVASGAGCLALGQNSSSSIEGSIALGSGSTSNRAITTGIRETSATSDGVVIGYNTTDRE 659

Qy        183 VIDAVNKA----GWRMKTTTANGQTGQ-----------------ADKFETVTSGTKVTF 220
              :: |::       :|  |  |:|  |                       |:    |  : :
Db        660 LLGALSLGTDGESYRQITNVADGSEAQDAVTVRQLQNAIGAVTTTPTKYYHANSTEEDSL 719
```

```
Qy        221 ASGNGTTATVSKDDQGNITVKYDVNVGDALNVNQLQN--SGWNLDSKAVAGSSGKVISGN 278
Db        720 AVGTDSLAMGAK-----TIVNADAGIGIGLNTLVMADAINGIAIGSNARANHANSIAMGN 774

Qy        279 VSP-----------------------SKGKMD---ETVNINAGN--------------N 297
Db        775 GSQTTRGAQTDYTAYNMDTPQNSVGEFSVGSEDGQRQITNVAAGSADTDAVNVGQLKVTD 834

Qy        298 IEITRNGKNI----------------------DIATSMTPQF---------------SSV 320
Db        835 AQVSRNTQSITNLNTQVSNLDTRVTNIENGIGDIVTTGSTKYFKTNTDGADANAQGADSV 894

Qy        321 SLGAGADAP--------TLSVDDE-GALNVGSKDANKPVRITNVAPGVKEGDVTNVAQLK 371
Db        895 AIGSGSIAAAENSVALGTNSVADEANTVSVGSSTQQR--RITNVAAGVNNTDAVNVAQLK 952

Qy        372 ------------------------------------------------GVAQ--------- 375
Db        953 ASEAGSVRYETNADGSVNYSVLNLGDGSGGTTRIGNVSAAVNDTDAVNYAQLKRSVEEAN 1012

Qy        376 --------NLNNRIDNVNGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYLGEAGYAI 427
Db       1013 TYTDQKMGEMNSKIKGIENKMSGGIASAMAMAGLPQAYAPGANMTSIAGGTFNGESAVAI 1072

Qy        428 GYSSISAGGNWIIKGTASGNSRGHFGASASVGYQW 462
Db       1073 GVSMVSESGGWVYKLQGTSNSQGDYSAAIGAGFQW 1107
```

RESULT 9
AH0110
probable surface protein (partial) YPO0902 [imported] - Yersinia pestis (strain CO92)
C;Species: Yersinia pestis
C;Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 02-Nov-2001
C;Accession: AH0110
R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,
Nature 413, 523-527, 2001
A;Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A;Reference number: AB0001; MUID:21470413; PMID:11586360
A;Accession: AH0110
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-658 <KUR>
A;Cross-references: GB:AL590842; PIDN:CAC89747.1; PID:g15978974; GSPDB:GN00175
C;Genetics:
A;Gene: YPO0902

Query Match . 13.5%; Score 317; DB 2; Length 658;
Best Local Similarity 23.6%; Pred. No. 1.4e-09;
Matches 117; Conservative 71; Mismatches 216; Indels 92; Gaps 15;

```
Qy          5 GLINVETEKLSFGANGKKVNIISDTKGLNFAKETAGTNGDTTVHLNGIGSTLTDMLLNTG 64
Db        217 GAVTSQANSIALGA--ASINTVGAQSSYSAYALTA---PQASVGELGIGTALGNRKITGV 271

Qy         65 ATTNVTNDNVT--------DDEKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDT 116
Db        272 AAGSASSDAVNVAQLTAVGDQVQQNTANITSLGGRVTTIEGSMASIANGGGVKYFHANST 331

Qy        117 -VEFLSADTKTTTVNVESKDNGKKT------EVKIGAKTSVIKEKDGKLVTGKGKGENGS 169
Db        332 QPDSVASGTNSVAIGPASLASGNAALASGAGAVAIG--DGAAASADGSVAIGQGSGDNGR 389

Qy        170 STDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTKVTFASGNGTTAT 229
Db        390 GVE---------------------------NYIGKYSNASNTSSGTVSVGNTATGETRT 421

Qy        230 VSKDDQGNITVKYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDET 289
Db        422 VSNVADG-------LQATDAVNLRQLDG----------IAASIVVVENNVSGLQNGTDGM 464
```

```
Qy        290 VNINAGNNI-EITRNGKNIDI--ATSMTPQFSSVSLGAGADAPTLS---------VDDEG 337
Db        465 FQVNNSSGLAKPSATGANSATGGAGSVASGNNSTAFGSGAKATAANSAALGANSVADRAN 524

Qy        338 ALNVGSKDANKPVRITNVAPGVKEGDVTNVAQLKGVAQN----LNNRIDNVNGNAR---- 389
Db        525 SVSVGSVGNER--QITNVAPATQGTDAVNFDQLKSISNQTNAYTNQRYSELKQDLRKQNS 582

Qy        390 ---AGIAQAIATAGLVQAYLPGKSMMAIGGGTYLGEAGYAIGYSSISAGGNWIIKGTASG 446
Db        583 VLSAGIASAMSMASLTQPYTSGSSMTTIGAASYRGQSALSLGVSSISDSGRWVSKLQASS 642

Qy        447 NSRGHFGASASVGYQW 462
Db        643 NTQGDFGIGVGVGYQW 658
```

RESULT 10
C82672
surface-exposed outer membrane protein XF1516 [imported] - Xylella fastidiosa (strain
C;Species: Xylella fastidiosa
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C;Accession: C82672
R;anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Seq
Nature 406, 151-157, 2000
A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A;Reference number: A82515; MUID:20365717; PMID:10910347
A;Note: for a complete list of authors see reference number A59328 below
A;Accession: C82672
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1004 <SIM>
A;Cross-references: GB:AE003981; GB:AE003849; NID:g9106543; PIDN:AAF84325.1; GSPDB:GN
A;Experimental source: strain 9a5c
R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer
as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Fr
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; La
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins
A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.
, F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri,
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawa
A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silv
M.; Tsuhako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.
A;Reference number: A59328
A;Contents: annotation
C;Genetics:
A;Gene: XF1516

Query Match 9.0%; Score 212.5; DB 2; Length 1004;
Best Local Similarity 23.1%; Pred. No. 0.0008;
Matches 130; Conservative 69; Mismatches 214; Indels 149; Gaps 26;

```
Qy          9 VETEKLSFGANGKKVNIISDTKGLNFAKETAGTNGDTTVHLNGIGSTLTDMLLNTGATTN 68
Db        156 VNSVALGAGSMASELNVISVGGG-------DGVTGPAVRRIVNVG----DGIGNNDAVNK 204

Qy         69 VTNDNVTDDEKKRAASVKDVL-----------NAGWNIKGVKPGTTA-SDNVDFVRTYD 115
Db        205 SQLDGVTASVNDVAASVKTIALTNQVTGSSVASASGKESTAIGSGAQAVADN-------- 256

Qy        116 TVEF-------------LSADT-----KTTTVNVESKD--------------NGKKTEV 142
Db        257 TVAFGGRAIANAVGASALGFDSHAKGINSTTVGTQSVSLGQGGVSLGYNSFVGEGSFNGL 316

Qy        143 KIGAKTSVIKEKDGKLVTGKG---KGENGSSTDEGEGL--VTAKEVI---------DAVN 188
Db        317 ALGSNSLVLLQGVDSVALGSGSMASEPNVVSVGSGDGLRGPAVRRIVNVGDGIGNNDAVN 376

Qy        189 KAGWRMKTTTANGQTGQADKFETV--TSGTKVTFASGNGTTAT-VSKDDQGNITVKYDVN 245
```

```
Db        377 KSQLDGVTASVNDVVASVKNIAGAIQITGSGVASVSGQDSTAAGASAQAAGDSSIA---- 432

Qy        246 VGDALNVNQLQNSGWNLD-------SKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNI 298
Db        433 LGARSRANAIGSSALGVDGHALGANSTALGGQSTAISEGGTSLG---YNSFVGQSATNGI 489

Qy        299 EITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGALNVGSKD-ANKPV--RITNV 355
Db        490 ALGSN--------AIVSGVNSVALGAGSVASELNV-----ISVGGGDGVTGPAVRRIVNV 536

Qy        356 APGVKEGDVTNVAQLKGVAQNLNN---RIDNVNGNAR---AGIAQAI---ATAGLVQAYL 406
Db        537 GDGIGNNDAVNKSQLDGVTASVNDVAASVKKIVGTIQITGSGVASAIGKDSTATGASAQA 596

Qy        407 PGKSMMAIG----------------------GGTYLGEAGYAIGYSSISAGGNWIIK--- 441
Db        597 VGDSSVALGTRATANAIGSSVLGVDSRARGINSTALGRQSNAIGDGSVSLGFNSFVRQSG 656

Qy        442 --GTASGNSRGHFGA-SASVGY 460
Db        657 EHGVALGTDAGVSGKDSIALGY 678
```

RESULT 11
AD0123
probable autotransporter protein yapH [imported] - Yersinia pestis (strain CO92)
C;Species: Yersinia pestis
C;Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 02-Nov-2001
C;Accession: AD0123
R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,
Nature 413, 523-527, 2001
A;Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A;Reference number: AB0001; MUID:21470413; PMID:11586360
A;Accession: AD0123
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-3705 <KUR>
A;Cross-references: GB:AL590842; PIDN:CAC89847.1; PID:g15979073; GSPDB:GN00175
C;Genetics:
A;Gene: yapH

Query Match 8.8%; Score 207; DB 2; Length 3705;
Best Local Similarity 24.4%; Pred. No. 0.0079;
Matches 129; Conservative 61; Mismatches 189; Indels 150; Gaps 27;

```
Qy          5 GLINVETEKLSFGANG-----KKVNIISD-TKGLNFAKETAG--TNGDTTVHLNGIGSTL 56
Db       2352 GTTTVTAGELILGANGAFGQTSLLNIASGASANINGYRQTVGAVTNTGTVTLGNGGELTS 2411

Qy         57 TDMLLNTGATTNVTNDNVTDDEKKRAASVKDVL--NAGWNIKGVKPGTTASDNVDFVRTY 114
Db       2412 TDTLINTG-MINVT-DGILNLENGGASSISGGLTGNGILNIKG----------GDF---- 2455

Qy        115 DTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKGKGENGSSTDEG 174
Db       2456 -------------TISIDNNGLAGQTNISDGA--SVTLGNGGTII---GTGNLGSSVIDV 2497

Qy        175 EG---LVTAKEVIDAVNKAGWRMKTTT----------------ANGQ--TGQADKF---- 209
Db       2498 LGDLNLVADNSLANVISGDGTINTTATVTLSGNSSFSGAHQIGTNGELTVGQASNLGASS 2557

Qy        210 ETVTSGTKVTFASGNGTTATVSKDDQGNITVKYDVNVGDALNVNQLQNSGWNLDSKAVAG 269
Db       2558 ATVNLGTLTSHLILNGVSESIANVLSGVAGSTVDI-IGGADTALTANNSGF-LGQYALAG 2615

Qy        270 SSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAG---- 325
Db       2616 NSKLTVA-----------STNNLGASSSVALAGTGDTLSL-SGFNGTFGNSVTGSGVLQV 2663

Qy        326 ----------------------ADAPTLSVDD----------EGALNVGSKDANKPVRIT 353
```

```
Db       2664 TDDAEVTLTSSNGVSNAVTIDIADA-TLNLDDIALFNHALTGNGLLNVAKNDASTAFDFG 2722

Qy        354 NVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVNGNARAGIAQA---IATAGLVQAYLPGKS 410
Db       2723 ATVGGAFTGTV----------NLNNSTFDLSGNNTTVLAQATLKLSSGNLTSV---GNG 2768

Qy        411 MMAIG-----GGTYL-----GEAGYAIGYSSISAGGNWIIKGTASGNSR 449
Db       2769 VQNIGTLAMNGGTLLFDNIVDNAGIITSDGTIAANS---INTTGGGEVR 2814
```

RESULT 12
AF0394
probable adhesin hmwA [imported] - Yersinia pestis (strain CO92)
C;Species: Yersinia pestis
C;Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 02-Nov-2001
C;Accession: AF0394
R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrel
Nature 413, 523-527, 2001
A;Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A;Reference number: AB0001; MUID:21470413; PMID:11586360
A;Accession: AF0394
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1910 <KUR>
A;Cross-references: GB:AL590842; PIDN:CAC92482.1; PID:g15981183; GSPDB:GN00175
C;Genetics:
A;Gene: hmwA

Query Match 8.5%; Score 200.5; DB 2; Length 1910;
Best Local Similarity 21.9%; Pred. No. 0.0077;
Matches 124; Conservative 71; Mismatches 191; Indels 179; Gaps 27;

```
Qy          6 LINVETEKLS--------FGAN------GKKVNIISDTKGLNFAKETAGTNG-------D 44
Db       1237 ILKGETEGVGATRKGIDFYGANTLNIIKGSQLSLLGENKG---AQDTAGGNGISYTSLAK 1293

Qy         45 TTVHLNGI----GSTLTDMLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKP 100
Db       1294 LTVNNNGSLKMEGRSTSGTGINFPSSNNTLVFNGDGDTLIKGSS---VAGTGAAISGVVN 1350

Qy        101 GTTASDNVDFVRTYDT-VEFLSADTKTTTVNVESKDNGKKTE---VKIGAKTSVIKEKDG 156
Db       1351 NSTGPMTIEGISTDGAGVHLFSAEHRIDRINV----TGSSTHAEGLRISGNAAIVD---- 1402

Qy        157 KLVTGKGKGENGSSTD----EGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETV 212
Db       1403 --TTLTGKSINGSGVKIDSLPGSSVVT-RSVLD----------NATLNGSSSSGKGVEIT 1449

Qy        213 TSGTKVTFASGNGTTATVSKDDQGNITVKYDVNVGDALNV---------- 252
Db       1450 SDINGIHHSSINGTTTGTG----------YGIDIGENSNVTGTSEADLLILQGVATTGTG 1499

Qy        253 --------NQLQNSGWNLDSKAVAG----------SSGKVISGNVSPSKG--------KM 286
Db       1500 TGIKLNGNNDLSNT--SLNSSAVDGIALDITGPLANQGNVILNGTASGSGIGAQVNGSLS 1557

Qy        287 DETVNINAGNNIEITRN----------------------------------GKNIDI 309
Db       1558 DSVVNGTSTNGIGVQINGSLKNSRINGISANGSGVKIDGESTLDNATLNGNSTEGKGVDL 1617

Qy        310 ATSMTPQFSSVSLGAGADAPTLSVDDEGALNVGSKDANKPVRITNVAPGVKEGDVTNVAQ 369
Db       1618 AANLSGNHGSVVHGDTVHGTGIDVGKDVTLSGGGTD--EPLTVSGNASGEKGTGV----Q 1671

Qy        370 LKGVAQNLNNRIDN--VNGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYLGEAGYAI 427
Db       1672 LGG-----NNTLDNTTLSGNATDGHGVEI------NSRLINNGNTTINGKT--SDDGHGV 1718

Qy        428 GYSSISAGGNWIIKGTASGNSRGHF 452
```

```
Db       1719 HINGAISGGE--ING-HSDNSHGVF 1740


RESULT 13
A43855
high-molecular-weight surface-exposed protein HMW1 - Haemophilus influenzae
C;Species: Haemophilus influenzae
C;Date: 10-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 08-Oct-1999
C;Accession: A43855
R;Barenkamp, S.J.; Leininger, E.
Infect. Immun. 60, 1302-1313, 1992
A;Title: Cloning, expression, and DNA sequence analysis of genes encoding nontypeable Ha
detella pertussis.
A;Reference number: A43855; MUID:92192797; PMID:1548058
A;Accession: A43855
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1536 <BAR>
A;Cross-references: GB:U08876; GB:M84616; NID:g475770; PIDN:AAA20527.1; PID:g475771
A;Note: sequence inconsistent with the nucleotide translation
A;Note: sequence extracted from NCBI backbone (NCBIN:89235, NCBIP:89239)

  Query Match             8.3%;  Score 195;  DB 2;  Length 1536;
  Best Local Similarity   24.4%;  Pred. No. 0.011;
  Matches  119;  Conservative   68;  Mismatches  191;  Indels  110;  Gaps   26;

Qy          8 NVETEKLSFGANGKKVNIISDTKGLNFAKETAGTN-----GDTTVHLNGIGSTLTDMLLN 62
              ||      |  : | : |  |   :|    ||| |      |: || |     :|:   |
Db        858 NVTLIGSDFDNHQKPLTIKKDVI-INSGNLTAGGNIVNIAGNLTVESNANFKAITNFTFN 916

Qy         63 TGAT-TNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLS 121
               |    |  | |::    |  |  ||: |:    | :   | :|       || |:   :
Db        917 VGGLFDNKGNSNIS--IAKGGARFKDIDNS----KNLSITTNSSS------TYRTIISGN 964

Qy        122 ADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKGK-------------GEN 167
                |   :|: ::  |  ||::||   |   :|:| |     |               |||
Db        965 ITNKNGDLNITNE--GSDTEMQIGGDVS---QKEGNLTISSDKINITKQITIKAGVDGEN 1019

Qy        168 GSSTDEGEGLVTAK----EVIDAVNKAGWRMKTTTAN-------GQTGQADKFETVTSGT 216
              |      :| |    ::   :| :|:     ||        | |  ||      |:
Db       1020 SDSDATNNANLTIKTKELKLTQDLNISGFNKAEITAKDGSDLTIGNTNSADG----TNAK 1075

Qy        217 KVTF---------ASGNGTTATVSKDDQGNITVKYDVNVGDALNVNQLQNSGWNLDSKAV 267
              ||||         | |:  |     :  |:     : |  |: :     |:|  :|:| |
Db       1076 KVTFNQVKDSKISADGHKVTLHSKVETSGS-----NNNTEDSSD----NNAGLTIDAKNV 1126

Qy        268 AGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIAT---SMTPQFSSVSLGA 324
                    :: |::  |      |:|:| :    |: |  |:  |    :| |  |:  |
Db       1127 ------TVNNNITSHK-----AVSISATSGEITTKTGTTINATTGNVEITAQTGSILGGI 1175

Qy        325 GADAPTLSVD-DEGALNVGSKDANKPVRITNVAPGVKEGDVTNVA--QLKGV-AQNLNNR 380
               : : ::::   |||| | :   |       |      | :| :|   :|| :   :::
Db       1176 ESSSGSVTLTATEGALAVSNISGN------TVTVTANSGALTTLAGSTIKGTESVTTSSQ 1229

Qy        381 IDNVNGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYLGEAGYAIGYSSISAGGNWII 440
               :: |    |  :  ||  |      :|   |   |  |||         || |   |
Db       1230 SGDIGGTISGGTVEVKATESLTT-----QSNSKIKATT--GEANVT------SATGT--I 1274

Qy        441 KGTASGNS 448
                || |||:
Db       1275 GGTISGNT 1282


RESULT 14
T31105
hypothetical protein 2 - Haemophilus ducreyi
C;Species: Haemophilus ducreyi
C;Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999
C;Accession: T31105
R;Ward, C.K.; Lumbley, S.R.; Latimer, J.L.; Cope, L.D.; Hansen, E.J.
J. Bacteriol. 180, 6013-6022, 1998
A;Title: Haemophilus ducreyi secretes a filamentous hemagglutinin-like protein.
A;Reference number: Z20984; MUID:99030326; PMID:9811662
A;Accession: T31105
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-4919 <WAR>
A;Cross-references: EMBL:AF057696; NID:g3929021; PID:g3929023; PIDN:AAC79761.1
C;Genetics:
A;Gene: lspA2

  Query Match             8.1%;  Score 190;  DB 2;  Length 4919;
  Best Local Similarity   22.5%;  Pred. No. 0.089;
  Matches  136;  Conservative   68;  Mismatches  213;  Indels  188;  Gaps   29;

Qy          7 INVETEKLSFGANGKKVNIISDTKGLNFAKETAGT---NGDTTVHLNGIGSTLTDMLLNT 63
              :|:   :|| ||   :|  ::    | :| |||    |    | ||   | |    |
Db        515 LNLNNSELS--ANNLTLNTSNNITLKNKSKFTAGNMTLNVTNNVTLNN-DSELAANNLTL 571

Qy         64 GATTNVT----------------NDNVTDDEKKRAAS-------VKDV------------ 88
                | |||                |||| : |   ::       ||:|
Db        572 NVTKNVTLNDASKLSANKLDLNVTDNVTLNSKSTLSAGELTFKKVKNVTLNNDSELAANN 631

Qy         89 --LNAGWNIKGVKPGTTASDNVDF----VRTYDTVEFL--SADTKTTTV----------- 129
                |||  |:         ::   |      :   || |    : |  :||:
Db        632 LSLNASHNVTLNNKSKLSAQKADIKAVNLTLNDTTELTAKNLDINSTTITNNGTIAGIFA 691

Qy        130 NVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKGK----------GENGSSTDEGEGLV 178
              |: ::    | :  | |: ::    :|     ||             :|   | |  ||
Db        692 NITTEKLNNKEKALILAEQNLNFTVNGSHYENKGDIVSKDKATVTFSKNSDFTSNGSKLV 751

Qy        179 TAKEVIDAVNKAGWRMKTTTANGQTGQADKFE---TVTSGTKVTFA-SGNGTTA-TVSKD 233
               |:           ::|    |    | |       ||     ||  ||| ||  |:
Db        752 NAQN----------QLKVNVNNFTISQGDDITLIGNVTLNASGTFTNSGNLTTVKTLDVG 801

Qy        234 DQGNITVKYDVNVGDALNV---NQLQNSG-------WNLDSKAVAGSSGKVI-------- 275
              |  | | | :: ||: |::    :: | |         |: |:|   ::| ::
Db        802 DIQNFTNKGNLTVGEDLHIKSKTKITNDGKLISIKNLNISSEADFINNGTLLGIEALKIA 861

Qy        276 -SGN----------------VSPSKGKMDETVN---INAGNNIEITRNGK--NIDIATSM 313
               ||                :| ::||  :| |    | :| |: ||   |   |:| ||
Db        862 TKGNFTNKEKAILASNSLLDISVAEGK--KTFNNGTIESGKNLNITNTGAFLNVDNATIR 919

Qy        314 TPQFSSVSLGAGADAPTLSVDDEGALNVGSK-DANKPVRITNVAPG----------VKE 361
                  | |       | :| : | |    : :       || : |            |:
Db        920 -------SFGVLNITSTGNVSNNGTLISNERLNITSAANFTNESNGTVMSNGLLNIIAKQ 972

Qy        362 GDVTNVAQLKGVAQ-NLNNRIDNVNGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYL 420
              |::||   :   | ||    ||:  :  : |:  ||   :             |  |
Db        973 GNITNKNLIASRQQLNLTAVADNITND--SNISNKIAVLHSL-------------GNISL 1017

Qy        421 GEAGYAIGYSSISAGGNWIIK-------------------GTAS-----GNSRGHFGAS 455
                      | || |  :|                     | ||      :: ||||
Db       1018 NSKDQVYNLGEIYAGNNISVKAHQLKNDVKLMGDITTKTKEGQASYKLYQASNGGHFGND 1077

Qy        456 ASVGY 460
                | ||
Db       1078 GSSGY 1082


RESULT 15
I40711
sapB protein - Campylobacter fetus
C;Species: Campylobacter fetus
C;Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 08-Oct-1999
C;Accession: I40711
R;Dworkin, J.; Tummuru, M.K.; Blaser, M.J.
J. Biol. Chem. 270, 15093-15101, 1995
A;Title: Segmental conservation of sapA sequences in type B Campylobacter fetus cells
A;Reference number: A56999; MUID:95318069; PMID:7797493
A;Accession: I40711
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-936 <RES>
A;Cross-references: EMBL:U25133; NID:g801999; PIDN:AAA79683.1; PID:g802000
C;Genetics:
```

A;Gene: sapB

Query Match 8.0%; Score 188.5; DB 2; Length 936;
Best Local Similarity 22.7%; Pred. No. 0.014;
Matches 119; Conservative 71; Mismatches 225; Indels 109; Gaps 25;

```
Qy          1 TDETGLINVETEK--LSFGANGK-KVNII-----------SDTKGLNFAKET---AGTNG 43
              || :  :||||     :| | |  ||:::           : :  |    :|      |
Db        310 TDASKSVNVETTGTITAFTAAGTGKVDVVAGKITALTADSATSVNLTATNDTITLTSANA 369

Qy         44 DTTVHLNGIGSTLTDMLLNTGATTNVTND-----NVTDDEKKRAASVKDVLNAGWNIKGV 98
               |:|:|   |:  :  : :  |  |:| |      ||        :||   |
Db        370 ATSVNLKTSGAAKSATITSANAAKNITIDATGVAAVTSATAVENLTVKHATNVTLAGNMD 429

Qy         99 KPGTTASDNVDF-----VRTYDTVEFLSADTKTTTVNVESKD-----NGKKTEVKIGA-- 146
              |  |   ||        :::  |:  :::      ::  :||     :|   :||:
Db        430 KLATVTLDNAALTAAIDIKSASTLNLINSSVNGHNISTAAKDVTVHLSGSAAKVKLNTTA 489

Qy        147 ---KTSVIKEK--DGKLVTGKGKGENGSSTDEGEG---LVTAKEVIDAVNKAGWRMKTTT 198
               :|  :|    |  |    |  :  | |  | |   ::   ||   ||     : ||
Db        490 ATDQTVTLKANATDNSLEFDSGTAKTTSVTASGSGKTLVIKGAEVETLVN-----IDTTA 544

Qy        199 AN-------GQTGQADKF--ETVTSGTKVTFASGNGTTATVSK--DDQGNITVKYDVNVG 247
              |       |:  |: ||  :| |   |: |    ||| |     |  || |:
Db        545 FNGAADVSFGKDAQSGKFSVKTGTGDDKIEFV---GTTLTEGSVIDGAGNDTIAMKSAAL 601

Qy        248 DALNVNQLQNSGWNL---DSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITR-- 302
              : |   ::|   |:   |: | |  |        :  :|   | |: ||    |  |
Db        602 TSANFTMIKNIE-NVAISDAVATADLSSSAFKNIIITTKEAADTTLTINKDQVINFTAAD 660

Qy        303 --------------NGKN--IDIATSMTPQFSSVSLGAGADAPTLSVDDE-GALNVGS-K 344
                           | |  : |       : :|::|| |    | :|     ||: |
Db        661 RGSVKLITVKLNDVTGANDVVKIVLDAAAKDASIALGTEATDKALVIDTGIETLNITSLV 720

Qy        345 DANKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVNGNARAGIAQAIATAGLVQA 404
              | |    |     |  |||::               ::|:|:  :  |  |||   :
Db        721 KATSPENTANTV-NAKLTDVTSII--------------IDGDAKITLGHA-GTAGTDYS 763

Qy        405 YLPGKSMMAIGGGTYLGEAGYAIGYSSISAGGNWIIKGTASGNS 448
              :      |:       :||     |:|: | |  |||  :  :|
Db        764 KVSMIDARAL-------KAGLTFDASAITLGANATIKGGSGADS 800
```

Search completed: October 6, 2003, 09:33:37
Job time : 16.2249 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: October 6, 2003, 09:06:20 ; Search time 7.13446 Seconds
(without alignments)
3045.266 Million cell updates/sec

Title: US-09-771-382-36
Perfect score: 2350
Sequence: 1 TDETGLINVETEKLSFGANG.....TASGNSRGHFGASASVGYQW 462

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	187.5	8.0	2249	1	OMPA_RICRI	P15921 rickettsia
2	187	8.0	2003	1	YDBA_ECOLI	P33666 escherichia
3	180.5	7.7	1655	1	OMPB_RICCN	Q9kka3 r outer mem
4	180	7.7	1654	1	OMPB_RICRI	Q53047 r outer mem
5	179.5	7.6	737	1	ALYS_ENTFA	P37710 enterococcu
6	178	7.6	1953	1	BIGA_SALTY	P25927 salmonella
7	177	7.5	1039	1	AG43_ECOLI	P39180 escherichia
8	177	7.5	1656	1	OMPB_RICJA	O06653 r outer mem
9	173.5	7.4	1577	1	HLYA_PROMI	P16466 proteus mir
10	173	7.4	1567	1	ICEN_XANCT	P18127 xanthomonas
11	172.5	7.3	933	1	SLAP_CAMFE	P35827 campylobact
12	170	7.2	1300	1	120K_RICRI	P14914 rickettsia
13	168	7.1	2021	1	OMPA_RICCN	Q52657 rickettsia
14	166	7.1	1645	1	OMPB_RICTY	P96989 r outer mem
15	165	7.0	1325	1	YDEK_ECOLI	P32051 escherichia
16	164	7.0	1025	1	SLAP_CAUCR	P35828 caulobacter
17	163	6.9	1569	1	YPJA_ECOLI	P52143 escherichia
18	161	6.9	1608	1	HLYA_SERMA	P15320 serratia ma
19	159	6.8	1286	1	AIDA_ECOLI	Q03155 escherichia
20	156.5	6.7	1007	1	Y741_CHLMU	Q9pjt6 chlamydia m
21	156	6.6	1861	1	APU_THETU	P38536 t amylopull
22	155	6.6	948	1	HPI1_DEIRA	P56867 deinococcus
23	154.5	6.6	1005	1	Y456_CHLTR	O84462 chlamydia t
24	152.5	6.5	918	1	YMJB_CAEEL	P34487 caenorhabdi
25	152	6.5	1694	1	IGA0_HAEIN	P44969 haemophilus
26	152	6.5	1702	1	IGA2_HAEIN	P45384 haemophilus
27	151	6.4	550	1	FLIC_SHIFL	Q08860 shigella fl
28	151	6.4	1643	1	OMPB_RICPR	Q53020 r outer mem
29	150.5	6.4	2660	1	YEEJ_ECO57	Q8x8v7 escherichia
30	150	6.4	1148	1	ICEK_PSESX	O30611 pseudomonas
31	150	6.4	1210	1	ICEN_PSEFL	P09815 pseudomonas
32	149.5	6.4	484	1	P60_LISMO	P21171 listeria mo
33	149.5	6.4	2358	1	YEEJ_ECOLI	P76347 escherichia
34	148.5	6.3	1150	1	APMU_PIG	P12021 sus scrofa
35	148.5	6.3	1196	1	ICEV_PSESX	O33479 pseudomonas
36	147	6.3	928	1	PM10_CHLPN	Q9rb65 chlamydia p
37	146.5	6.2	3178	1	YS89_CAEEL	Q09624 caenorhabdi
38	146	6.2	2334	1	WAPA_BACSU	Q07833 bacillus su
39	144.5	6.1	504	1	FLIC_SALBU	Q06969 salmonella
40	144.5	6.1	1153	1	PVDB_PLAKN	P50493 plasmodium
41	144	6.1	585	1	KPYK_BACLI	P51181 bacillus li
42	144	6.1	678	1	YF48_MYCTU	Q10778 mycobacteri
43	144	6.1	3591	1	FHAB_BORPE	P12255 bordetella
44	143	6.1	537	1	TEE6_STRPY	P18481 streptococc
45	143	6.1	1200	1	ICEN_PSESY	P06620 pseudomonas

ALIGNMENTS

RESULT 1
OMPA_RICRI
ID OMPA_RICRI STANDARD; PRT; 2249 AA.
AC P15921;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Outer membrane protein A precursor (190 kDa antigen) (Cell surface
DE antigen) (rOmpA) (rOmp A).
GN OMPA.
OS Rickettsia rickettsii.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Rickettsiaceae; Rickettsieae; Rickettsia.
OX NCBI_TaxID=783;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=R;
RX MEDLINE=90354033; PubMed=2117568;
RA Anderson B.E., McDonald G.A., Jones D.C., Regnery R.L.;
RT "A protective protein antigen of Rickettsia rickettsii has tandemly
RT repeated, near-identical sequences.";
RL Infect. Immun. 58:2760-2769(1990).
CC -!- FUNCTION: ELICITS PROTECTIVE IMMUNITY.
CC -!- SUBCELLULAR LOCATION: CELL WALL. THIS BACTERIUM IS COVERED BY A
CC S-LAYER WITH HEXAGONAL SYMMETRY.
CC -!- PTM: GLYCOSYLATED (PROBABLE).
CC -!- SIMILARITY: BELONGS TO THE RICKETTSIAE OMPA/OMPB FAMILY.
CC ---

CC ---
DR EMBL; M31227; AAA26380.1; -.
DR PIR; A41477; A41477.
DR InterPro; IPR006315; Autotransport.
DR InterPro; IPR005546; Autotransporter.
DR Pfam; PF03797; Autotransporter; 1.
DR TIGRFAMs; TIGR01414; autotrans_barl; 3.
KW Antigen; Repeat; Signal; Cell wall; S-layer; Glycoprotein.
FT SIGNAL 1 28 POTENTIAL.
FT CHAIN 29 2249 OUTER MEMBRANE PROTEIN A.
FT DOMAIN 212 1180 13 X APPROXIMATE TANDEM REPEATS.
FT REPEAT 212 286 A (TYPE I).
FT REPEAT 287 358 B (TYPE II).
FT REPEAT 359 430 C (TYPE II).
FT REPEAT 431 505 D (TYPE I).
FT REPEAT 506 577 E (TYPE II).
FT REPEAT 578 652 F (TYPE I).
FT REPEAT 653 724 G (TYPE II).
FT REPEAT 725 799 H (TYPE I).
FT REPEAT 800 874 I (TYPE I).
FT REPEAT 875 949 J (TYPE I).

```
FT   REPEAT      950   1021       K (TYPE II).
FT   REPEAT     1022   1093       L (TYPE II).
FT   REPEAT     1094   1165       M (TYPE II).
FT   REPEAT     1166   1180       TYPE I (INCOMPLETE).
SQ   SEQUENCE   2249 AA;  224333 MW;  A9D6646C089DF087 CRC64;

  Query Match             8.0%;  Score 187.5;  DB 1;  Length 2249;
  Best Local Similarity   22.4%;  Pred. No. 0.034;
  Matches  124;  Conservative   60;  Mismatches  214;  Indels  155;  Gaps   27;

Qy          6 LINVETEKLSFGANGKKVNIISDTKGLNFAKETAGTNGDT--TVHLNGIGSTLTDMLLNT 63
              :|   | ||:   |   |  :::   :        | ||    ::|||  | :|  : ||
Db        822 VIKATTTKLT---NAASVLTLTNANAVLTGAVDNTTGGDNVGVLNLNGALSQVTGDIGNT 878

Qy         64 ----------------GATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPGTTASDN 107
                            ||    |   :|:     ||||  : ||   : |    ||  ||
Db        879 NSLATISVGAGTATLGGAVIKATTTKLTN-----AASVLTLTNANAVLTGAIDNTTGGDN 933

Qy        108 VDFVRTYDTVEFLSADTKTT----TVNVESKD---NG---KKTEVKIGAKTSVIKEKDGK 157
              |  :     :  :: |   |    |    |::| :       |  | |  |:    | :|  :
Db        934 VGVLNLNGALSQVTGDIGNTNSLATISVGAGTATLGGAVIKATTTKLTDAASAVKFTNPV 993

Qy        158 LVTG----KGKGENGSSTDEGEGLVTAK----EVIDAVNKA-------GWRMKTTTAN-- 200
              :|||     |   ||  |  |   ||        :  ||         |  :|  | |
Db        994 VVTGAIDNTGNANNGIVTFTGNSTVTGNVGNTNALATVNVGAGLLQVQGGVVKANTINLT 1053

Qy        201 ------------GQTGQADKFETVTSGTKVTFASGNGTTATVSKDDQGNITVKYDVNVGD 248
                         ||  |      :|  |||   :  |  |     ||      ||||
Db       1054 DNASAVTFTNPVVVTGAIDNTGNANNGI-VTFTGNSTVTGNV-----GNTNALATVNVGA 1107

Qy        249 AL-----------NVN-----------------QLQNSGWNLDSKAVAGSSGKVISGNV 279
              |           :|                 : |:| | ::  |  :    ::|::
Db       1108 GLLQVQGGVVKANTINLTDNASAVTFTNPVVVTGAIDNTG-NANNGIVTFTGNSTVTGDI 1166

Qy        280 SPSKGKMDETVNINAGNNIEITRNG----KNIDIATSMTPQFSSVSLGAGADAP-----T 330
                |  :    :| |  | :   |     |||     | :|:    | |   |
Db       1167 ----GNTNALATVNVGAGITLQAGGSLAANNIDFGARSTLEFNGPLDGGGKAIPYYFKGA 1222

Qy        331 LSVDDEGALNVGSKDANK------PVRITNVAPG------VKEGDVTNVAQLKGVAQNLN 378
              ::  :   ||| :|           |   |:  |          |||| :      |||:|
Db       1223 IANGNNAILNVNTKLLTASHLTIGTVAEINIGAGNLFTIDASVGDVTILN-----AQNIN 1277

Qy        379 NR-------IDNVNGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYLGEAGYAIGYSS 431
              |       : |: |    |:   :  | ||    ||                |  :  | :
Db       1278 FRARDSVLVLSNLTG---VGVNNILLAADLV---APGADE---------GTVVFNGGVNG 1322

Qy        432 ISAGGNWIIKGTA 444
              :: | |  : |||
Db       1323 LNVGSN--VAGTA 1333


RESULT 2
YDBA_ECOLI
ID   YDBA_ECOLI     STANDARD;      PRT;  2003 AA.
AC   P33666; P76087; P76088; P76856; P76857; P76859;
DT   01-FEB-1994 (Rel. 28, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Hypothetical protein ydbA.
GN   YDBA OR B1401/B1405.
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12 / MG1655;
RX   MEDLINE=97426617; PubMed=9278503;
RA   Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA   Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA   Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA   Mau B., Shao Y.;
RT   "The complete genome sequence of Escherichia coli K-12.";
RL   Science 277:1453-1474(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12;
RX   MEDLINE=97251357; PubMed=9097039;
RA   Aiba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K., Itoh T.,
RA   Kasai H., Kashimoto K., Kimura S., Kitakawa M., Kitagawa M.,
RA   Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,
RA   Nakade S., Nakamura Y., Nashimoto H., Nishio Y., Oshima T., Saito N.,
RA   Sampei G., Seki Y., Sivasundaram S., Tagami H., Takeda J.,
RA   Takemoto K., Takeuchi Y., Wada C., Yamamoto Y., Horiuchi T.;
RT   "A 570-kb DNA sequence of the Escherichia coli K-12 genome
RT   corresponding to the 28.0-40.1 min region on the linkage map.";
RL   DNA Res. 3:363-377(1996).
RN   [3]
RP   SEQUENCE OF 464-2003 FROM N.A.
RC   STRAIN=K12;
RX   MEDLINE=92190338; PubMed=1665988;
RA   Moszer I., Glaser P., Danchin A.;
RT   "Multiple IS insertion sequences near the replication terminus in
RT   Escherichia coli K-12.";
RL   Biochimie 73:1361-1374(1991).
CC   -!- SIMILARITY: TO S.TYPHIMURIUM ORF NEAR CYSG (AC P25928).
CC   -!- CAUTION: THIS IS A CONCEPTUAL TRANSLATION; THE GENE CODING FOR
CC       THIS PROTEIN IS INTERRUPTED BY A HYBRID IS2D/IS30C ELEMENT
CC       BETWEEN AMINO ACIDS 839 AND 840.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AE000237; AAC74483.1; ALT_SEQ.
DR   EMBL; AE000237; AAC74487.1; ALT_SEQ.
DR   EMBL; D90778; BAA15009.1; ALT_SEQ.
DR   EMBL; D90778; BAA18880.1; ALT_SEQ.
DR   EMBL; D90779; BAA18881.1; ALT_SEQ.
DR   EMBL; X62680; -; NOT_ANNOTATED_CDS.
DR   EcoGene; EG11307; ydbA.
KW   Hypothetical protein; Complete proteome.
FT   CONFLICT    489    489       I -> V (IN REF. 2).
FT   CONFLICT    495    495       I -> V (IN REF. 2).
SQ   SEQUENCE   2003 AA;  205949 MW;  B83A12C8B53220EE CRC64;

  Query Match             8.0%;  Score 187;  DB 1;  Length 2003;
  Best Local Similarity   25.0%;  Pred. No. 0.032;
  Matches  137;  Conservative   73;  Mismatches  207;  Indels  130;  Gaps   33;

Qy          2 DETGLINV-ETEKLSFGANGKKVNIISD-----TKG-----LNFAKETAGTNGDTTVHLN 50
              |  | : | : | :    :| |  : ::     | |      :|    ||  || |||  :
Db        276 DNKGTMTVTDPESMGIQIDGDKAIVNNEGESTITNGGTGTQINGDDATANNNGKTTV--D 333

Qy         51 GIGSTLTDMLLNTGAT-------------TNVTNDNVTDDEKKRAASVKDVLNAGWNIK 96
              |  || |::  | |                  ::| |: | |  |   :| |  : |  :
Db        334 GKDSTGTEINGNNGKVIQDGDLDVSGGGHGIDITGDSATVD-NKGTMTVTDPESIGIQVD 392

Qy         97 GVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKK---TEVKIGAKTSVIKE 153
              |       ::    :    |   :: | |   | :: :||    ||:  |    ||
Db        393 G-DQAVVNNEGESAITNGGTGTQINGDDATANNNGKTTVDGKDSTGTEI-AGNNGKVI-- 448

Qy        154 KDGKL-VTGKGKG---ENGSSTDEGEGLVTAKE------VIDA----VNKAGWRMKTTTA 199
              :|| | |:| | |      |:| : :| :|  :        ||     ||  |   ::|
Db        449 QDGDLDVSGGGHGIDITGDSATVDNKGTMTVTDPESIGIQIDGDQAIVNNEG---ESTIT 505

Qy        200 NGQTGQADKFETVTSGTKVTFASGNGTTATVSKDD-----QGNITVKYDVNVGDALNV-- 252
              || ||      |  :|   | |: :| |    ||       |||  :   ||:  :| |
Db        506 NGGTG------TQINGNDAT-ANNSGKTTVDGKDSTGTKIAGNIGI---VNLDGSLTVTG 555

Qy        253 --NQLQNSGWN--LDSK-----AVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRN 303
                : ::| | |  :::|     :  || | :|:|  :      |  |: |      || |
```

```
Db        556 GAHGVENIGDNGTVNNKGDIVVSDTGSIGVLINGEGATVSNTGDVNVS-NEATGFSITTN 614

Qy        304 GKNIDIATSM-TPQFSS-VSLGAGADAPTLSVDD-------EGALNVGSKDANKPVRITN 354
              : :| ||    ||: | |    :: ||:  |          :|| | |||  | ||
Db        615 SGKVSLAGSMQVGDFSTGVDLNGNNNSVTLAAKDLKVVGQKATGINV-SGDAN-TVNITG 672

Qy        355 VAPGVKEGDVTNVAQL---KGVAQNLNNRIDNVNGNARAGI---------------AQAI 396
                :    | |:      |  |:    :||  : :  :                   |
Db        673 NVLVDKDKTADNAAEYFFDPSVGINVYGSDNNVTLDGKLTVVSDSEVTSRQSNLFDGSAE 732

Qy        397 ATAGLVQAYLPGKSMMAIGGGTYLGE--------------AGYAI-------GYSSISAG 435
                |:|||     | ::   ||   :||               ||:        | ||:
Db        733 KTSGLV-VIGDGNTVNMNGGLELIGEKNALADGSQVTSLRTGYSYTSVIVVSGESSVYLN 791

Qy        436 GNWIIKG 442
                |:  | |
Db        792 GDTTISG 798


RESULT 3
OMPB_RICCN
ID   OMPB_RICCN     STANDARD;      PRT;  1655 AA.
AC   Q9KKA3; Q9KK98; Q9XC45;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Outer membrane protein B precursor (168 kDa surface-layer protein)
DE   (Surface protein antigen) (Cell surface antigen 5) (Sca5) (rOmpB)
DE   (rOmp B) [Contains: 120 kDa surface-exposed protein (Surface protein
DE   antigen) (120 kDa outer membrane protein ompB); 32 kDa beta peptide].
GN   OMPB OR RC1085.
OS   Rickettsia conorii.
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC   Rickettsiaceae; Rickettsieae; Rickettsia.
OX   NCBI_TaxID=781;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Malish 7;
RX   MEDLINE=21442074; PubMed=11557893;
RA   Ogata H., Audic S., Renesto-Audiffren P., Fournier P.-E., Barbe V.,
RA   Samson D., Roux V., Cossart P., Weissenbach J., Claverie J.-M.,
RA   Raoult D.;
RT   "Mechanisms of evolution in Rickettsia conorii and R. prowazekii.";
RL   Science 293:2093-2098(2001).
RN   [2]
RP   SEQUENCE OF 33-1649 FROM N.A.
RC   STRAIN=Indian tick typhus, and Malish 7;
RX   MEDLINE=20393643; PubMed=10939649;
RA   Roux V., Raoult D.;
RT   "Phylogenetic analysis of members of the genus Rickettsia using the
RT   gene coding the outer-membrane protein rOmpB (ompB).";
RL   Int. J. Syst. Evol. Microbiol. 50:1449-1455(2000).
RN   [3]
RP   SEQUENCE OF 353-1655 FROM N.A.
RC   STRAIN=Malish 7;
RA   Stenos J., Walker D.;
RT   "The rickettsial outer membrane protein A and B genes of Rickettsia
RT   australis, the most divergent rickettsia of the spotted fever group.";
RL   Submitted (MAY-1999) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: THE 120 KDA SURFACE-EXPOSED PROTEIN IS A MAJOR
CC       STRUCTURAL PROTEIN WHICH MAY PLAY A ROLE AS A RICKETTSIAL
CC       VIRULENCE FACTOR AND/OR IMMUNOGEN DURING INFECTION (BY
CC       SIMILARITY).
CC   -!- FUNCTION: THE 32 KDA BETA PEPTIDE MAY SERVE AS A MEMBRANE ANCHOR
CC       (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: CELL WALL. THIS BACTERIUM IS COVERED BY A S-
CC       LAYER WITH HEXAGONAL SYMMETRY (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE RICKETTSIAE OMPA/OMPB FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE008659; AAL03623.1; -.
DR   EMBL; AF123721; AAF34124.1; -.
DR   EMBL; AF123726; AAF34129.1; -.
DR   EMBL; AF149110; AAD39533.1; -.
DR   PIR; E97835; E97835.
DR   InterPro; IPR006315; Autotransport.
DR   InterPro; IPR005546; Autotransporter.
DR   Pfam; PF03797; Autotransporter; 1.
DR   TIGRFAMs; TIGR01414; autotrans_barl; 2.
KW   Antigen; S-layer; Cell wall; Complete proteome.
FT   CHAIN         1   1334       120 kDa SURFACE-EXPOSED PROTEIN.
FT   CHAIN      1335   1655       32 kDa BETA PEPTIDE.
FT   VARIANT      61     61       P -> A (IN STRAIN INDIAN TICK TYPHUS).
FT   VARIANT      75     75       G -> S (IN STRAIN INDIAN TICK TYPHUS).
FT   VARIANT      78     78       K -> N (IN STRAIN INDIAN TICK TYPHUS).
FT   VARIANT     251    251       V -> A (IN STRAIN INDIAN TICK TYPHUS).
FT   VARIANT     413    413       N -> D (IN STRAIN INDIAN TICK TYPHUS).
FT   VARIANT     959    959       I -> V (IN STRAIN INDIAN TICK TYPHUS).
FT   VARIANT     988    988       A -> T (IN STRAIN INDIAN TICK TYPHUS).
FT   VARIANT    1139   1139       R -> L (IN STRAIN INDIAN TICK TYPHUS).
FT   CONFLICT    353    354       KD -> GH (IN REF. 3).
FT   CONFLICT    776    776       F -> S (IN REF. 3).
FT   CONFLICT   1159   1159       E -> D (IN REF. 3).
FT   CONFLICT   1177   1177       G -> S (IN REF. 3).
FT   CONFLICT   1492   1492       H -> R (IN REF. 3).
SQ   SEQUENCE   1655 AA;  168342 MW;  E49E19377D5FCE37 CRC64;

  Query Match             7.7%;  Score 180.5;  DB 1;  Length 1655;
  Best Local Similarity   24.2%;  Pred. No. 0.054;
  Matches  121;  Conservative   59;  Mismatches  225;  Indels   95;  Gaps   26;

Qy         19 NGKKVNIISDTKGLNFAKETAGTNGDTTV-HLNGIGSTLTDMLLNTGATTNVT-NDNV-T 75
              |||   |      :  ::|   |     |: :|:  |       :   :| | |  |
Db        340 NGKVATIDGQVYAKDMVIQSANAVGQVNFRHIVDVGTDGTTAFKTAASKVAITQNSNFGT 399

Qy         76 DDEKKRAASV----KDVLNAGWNIKGVKPGTTASDNVDFVRTYD---TVEFLSADTKTTT 128
               |   || :       ||  :      || ||      | |:|   |:   |||
Db        400 TDFGNLAAQIIVPNTMTLNGNFTGDASNPGNTAG-----VITFDANGTLASASADANVAV 454

Qy        129 VN----VESKDNG-------KKTEVKIGAKTSVIKEKDGKLVTGK--------GKGENGS 169
               |    :|:   |           |:::|   || |  || :: ||        |    |:
Db        455 TNNITAIEASGAGVVQLSGTHAAELRLGNAGSVFKLADGTVINGKVNQTALVGGALAAGT 514

Qy        170 STDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVT---------SGTKVTF 220
               | :|   :|     | |  |          | |   |  :|:|          :|  : |
Db        515 ITLDGSATITG----DIGNAGG----AAALQGITLANDATKTLTLGGANIIGANGGTINF 566

Qy        221 ASGNGTTATVSKDDQGNITVKYDV-------NVGDALNVNQLQNSGWNLDSKAVAGSSGK 273
               :  ||    |   | || | :|:         | || ::   |    |      |  ::
Db        567 QANGGTIKLTS--TQNNIVVDFDLAIATDQTGVVDASSLTNAQTLTINGKIGTVGANNKT 624

Qy        274 VISGNVSPSKGKM-DETVNIN---AGNN--IEITRNGKNIDIATSMTPQ----FSSV--- 320
               :   |: ||  : |  | ||    |||  ::   |   |   |:  |    |: |
Db        625 LGQFNIGSSKTVLSDGDVAINELVIGNNGAVQFAHNTYLITRTTNAAGQGKIIFNPVVNN 684

Qy        321 --SLGAGADAPTLSVDDEGALNVGSKDANKPVRITNVAPGVKEGDVTNVAQLKGVAQNLN 378
                 :|  | :  : : :    :| ||| |     : ||  ||     ||:        |:
Db        685 NTTLATGTNLGS-ATNPLAEINFGSKGAANVDTVLNVGKGVNL-YATNITTTDA---NVG 739

Qy        379 NRIDNVNG-NARAGIAQAIATAGLVQ-------AYLPGKSMMAIGGGTYLGEAGYAIGYS 430
               : | |  | |  :|       | |  |       |   | :: :|  |: |    |  |
Db        740 SFIFNAGGTNIVSG------TVGGQQGNKFNTVALDNGTTVKFLGNATFNGNTTIAAN-S 792

Qy        431 SISAGGNWIIKGTASGNSRG 450
               ::  |||:     || :  |
Db        793 TLQIGGNYTADFVASADGTG 812


RESULT 4
```

```
OMPB_RICRI
ID   OMPB_RICRI     STANDARD;      PRT;  1654 AA.
AC   Q53047;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Outer membrane protein B precursor (168 kDa surface-layer protein)
DE   (Surface protein antigen) (Cell surface antigen 5) (Sca5) (rOmpB)
DE   (rOmp B) [Contains: 120 kDa surface-exposed protein (Surface protein
DE   antigen) (120 kDa outer membrane protein ompB); 32 kDa beta peptide].
GN   OMPB.
OS   Rickettsia rickettsii.
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC   Rickettsiaceae; Rickettsieae; Rickettsia.
OX   NCBI_TaxID=783;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=R;
RX   MEDLINE=92167802; PubMed=1724278;
RA   Gilmore R.D. Jr., Cieplak W. Jr., Policastro P.F., Hackstadt T.;
RT   "The 120 kilodalton outer membrane protein (rOmp B) of Rickettsia
RT   rickettsii is encoded by an unusually long open reading frame:
RT   evidence for protein processing from a large precursor.";
RL   Mol. Microbiol. 5:2361-2370(1991).
RN   [2]
RP   SEQUENCE OF 279-1654 FROM N.A.
RC   STRAIN=R;
RX   MEDLINE=90136087; PubMed=2515418;
RA   Gilmore R.D. Jr., Joste N., McDonald G.A.;
RT   "Cloning, expression and sequence analysis of the gene encoding the
RT   120 kD surface-exposed protein of Rickettsia rickettsii.";
RL   Mol. Microbiol. 3:1579-1586(1989).
CC   -!- FUNCTION: THE 120 kDa SURFACE-EXPOSED PROTEIN IS A MAJOR
CC       STRUCTURAL PROTEIN WHICH MAY PLAY A ROLE AS A RICKETTSIAL
CC       VIRULENCE FACTOR AND/OR IMMUNOGEN DURING INFECTION.
CC   -!- SUBCELLULAR LOCATION: CELL WALL. THIS BACTERIUM IS COVERED BY A S-
CC       LAYER WITH HEXAGONAL SYMMETRY.
CC   -!- SIMILARITY: BELONGS TO THE RICKETTSIAE OMPA/OMPB FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X16353; CAA34403.1; -.
DR   PIR; S18227; S18227.
DR   InterPro; IPR006315; Autotransport.
DR   InterPro; IPR005546; Autotransporter.
DR   Pfam; PF03797; Autotransporter; 1.
DR   TIGRFAMs; TIGR01414; autotrans_barl; 2.
KW   Antigen; S-layer; Cell wall.
FT   CHAIN         1   1333       120 kDa SURFACE-EXPOSED PROTEIN.
FT   CHAIN      1334   1654       32 kDa BETA PEPTIDE.
FT   DOMAIN     1181   1188       POLY-THR.
SQ   SEQUENCE   1654 AA;  168184 MW;  D7AB70FB7087F618 CRC64;

  Query Match             7.7%;  Score 180;  DB 1;  Length 1654;
  Best Local Similarity   23.9%;  Pred. No. 0.058;
  Matches  123;  Conservative   64;  Mismatches  203;  Indels  124;  Gaps   29;

Qy         19 NGKKVNIISDTKGLNFAKETAGTNGDTTV-HLNGIGSTLTDMLLNTGATTNVTNDN--VT 75
              |||   |      :   ::|   |     |: :|:  |       :   :| |:
Db        340 NGKVATIDGQVYAKDMVIQSANATGQVNFRHIVDVGADGTTAFKTAASKVTITQDSNFGN 399

Qy         76 DDEKKRAASVK--DVLNAGWNIKG--VKPGTTASDNVDFVRTYD---TVEFLSADTKTTT 128
                 |    || :|  : :    |  |    || ||      | |:|   |:|  |||
Db        400 TDFGNLAAQIKVPNAITLTGNFTGDASNPGNTAG-----VITFDANGTLESASADANVAV 454

Qy        129 VN----VESKDNG-------KKTEVKIGAKTSVIKEKDGKLVTGK--------GKGENGS 169
                  |    :|:   |          |:::|   |: |  || :: ||        |    |:

Db        455 TNNITAIEASGAGVVQLSGTHAAELRLGNAGSIFKLADGTVINGKVNQTALVGGALAAGT 514

Qy        170 STDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTKVTFASGNGT--- 226
                | :|   :|    |      |    : | ||      |  :|:| |       :| ||
Db        515 ITLDGSATITGD--IGNAGGAAALQRITLAN------DAKKTLTLGGANIIGAGGGTIDL 566

Qy        227 ---------TATVSKDDQGNITVKYDV-------NVGDALNVNQLQNSGWN-------LD 263
                      |:|     | || | :|:        | || ::   |    |        :
Db        567 QANGGTIKLTST-----QNNIVVDFDLAIATDQTGVVDASSLTNAQTLTINGKIGTIGAN 621

Qy        264 SKAVA----GSSGKVIS-GNVSPSKGKMDETVNINAG------NNIEITR------NGKN 306
              :| :     |||  |:| |||:     ::| |  | |       :   |||       ||
Db        622 NKTLGQFNIGSSKTVLSNGNVA-----INELVIGNDGAVQFAHDTYLITRTTNAAGQGKI 676

Qy        307 I--DIATSMTPQFSSVSLGAGADAPTLSVDDEGALNVGSKDANKPVRITNVAPGVKEGDV 364
              |   :  : |   :  :||:       : :    :| |||  |    : ||  ||
Db        677 IFNPVVNNGTTLAAGTNLGS-------ATNPLAEINFGSKGVNVDT-VLNVGEGVNL-YA 727

Qy        365 TNVAQLKGVAQNLNNRIDNVNG-NARAGIAQAIATAGLVQ-------AYLPGKSMMAIGG 416
              ||:        |: : : |  | |  :|      | |  |       |   | ::  :|
Db        728 TNITTTDA---NVGSFVFNAGGTNIVSG------TVGGQQGNKFNTVALENGTTVKFLGN 778

Qy        417 GTYLGEAGYAIGYSSISAGGNWIIKGTASGNSRG 450
              |: |    |   |:: |||:     || :  |
Db        779 ATFNGNTTIAAN-STLQIGGNYTADCVASADGTG 811


RESULT 5
ALYS_ENTFA
ID   ALYS_ENTFA     STANDARD;      PRT;   737 AA.
AC   P37710;
DT   01-OCT-1994 (Rel. 30, Created)
DT   15-SEP-2003 (Rel. 42, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Autolysin precursor (EC 3.2.1.-) (Peptidoglycan hydrolase)
DE   (Beta-glycosidase).
GN   EF0799.
OS   Enterococcus faecalis (Streptococcus faecalis).
OC   Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.
OX   NCBI_TaxID=1351;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91358349; PubMed=1679432;
RA   Beliveau C., Potvin C., Trudel J., Asselin A., Bellemare G.;
RT   "Cloning, sequencing, and expression in Escherichia coli of a
RT   Streptococcus faecalis autolysin.";
RL   J. Bacteriol. 173:5619-5623(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=V583 / ATCC 700802;
RX   MEDLINE=22550857; PubMed=12663927;
RA   Paulsen I.T., Banerjei L., Myers G.S.A., Nelson K.E., Seshadri R.,
RA   Read T.D., Fouts D.E., Eisen J.A., Gill S.R., Heidelberg J.F.,
RA   Tettelin H., Dodson R.J., Umayam L., Brinkac L., Beanan M.,
RA   Daugherty S., DeBoy R.T., Durkin S., Kolonay J., Madupu R., Nelson W.,
RA   Vamathevan J., Tran B., Upton J., Hansen T., Shetty J., Khouri H.,
RA   Utterback T., Radune D., Ketchum K.A., Dougherty B.A., Fraser C.M.;
RT   "Role of mobile DNA in the evolution of vancomycin-resistant
RT   Enterococcus faecalis.";
RL   Science 299:2071-2074(2003).
CC   -!- FUNCTION: Hydrolyzes the cell wall of E.faecalis and
CC       M.lysodeikticus. May play an important role in cell wall growth
CC       and cell separation.
CC   -!- SUBCELLULAR LOCATION: Secreted (Probable).
CC   -!- DOMAIN: LYSM REPEATS ARE THOUGHT TO BE INVOLVED IN PEPTIDOGLYCAN
CC       BINDING.
CC   -!- SIMILARITY: BELONGS TO FAMILY 73 OF GLYCOSYL HYDROLASES.
CC   -!- SIMILARITY: Contains 6 LysM repeats.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
```

```
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; M58002; AAA67325.1; -.
DR   EMBL; AE016949; AAO80613.1; -.
DR   PIR; A38109; A38109.
DR   TIGR; EF0799; -.
DR   InterPro; IPR002901; Amidase_4.
DR   InterPro; IPR002482; LysM.
DR   Pfam; PF01832; Amidase_4; 1.
DR   Pfam; PF01476; LysM; 5.
DR   SMART; SM00257; LysM; 6.
DR   SMART; SM00047; LYZ2; 1.
KW   Hydrolase; Glycosidase; Bacteriolytic enzyme; Cell wall;
KW   Cell division; Septation; Repeat; Signal.
FT   SIGNAL        1     53       POTENTIAL.
FT   CHAIN        54    737       AUTOLYSIN.
FT   REPEAT      363    405       LYSM 1.
FT   REPEAT      431    473       LYSM 2.
FT   REPEAT      499    541       LYSM 3.
FT   REPEAT      567    609       LYSM 4.
FT   REPEAT      633    675       LYSM 5.
FT   REPEAT      695    737       LYSM 6.
FT   CONFLICT     85     85       T -> I (IN REF. 1).
FT   CONFLICT    118    118       A -> V (IN REF. 1).
FT   CONFLICT    143    143       A -> T (IN REF. 1).
FT   CONFLICT    417    417       S -> N (IN REF. 1).
FT   CONFLICT    449    449       S -> T (IN REF. 1).
FT   CONFLICT    476    476       A -> T (IN REF. 1).
FT   CONFLICT    484    484       N -> S (IN REF. 1).
FT   CONFLICT    567    632       MISSING (IN REF. 1).
SQ   SEQUENCE   737 AA;  77025 MW;  ABB16BD506AC7507 CRC64;

  Query Match             7.6%;  Score 179.5;  DB 1;  Length 737;
  Best Local Similarity   23.0%;  Pred. No. 0.024;
  Matches  111;  Conservative   55;  Mismatches  208;  Indels  109;  Gaps   21;

Qy         15 SFGANGKKVNIISDTKGLNFA----KETAGTNGDTTVHLNGIGSTLTDMLLNTGATTNVT 70
              ||  |   :   |   |: :     |    :  | |  |   |   ||   |      :|
Db        268 SFQDNAHVLKTTSFQAGVYYYAGAWKSNTSSYRDATAWLT--GRYATDPSYNAKLNNVIT 325

Qy         71 NDNVTDDEKKRAASVKDVLNAGWNIKG--VKPGTTASDNVDFVRTY------DTVEFLSA 122
                |:|           |  ::| |  |  | |||  |:|     ||      ||:  ::|
Db        326 AYNLTQ---------YDTPSSGGNTGGGTVNPGTGGSNNQSGTNTYYTVKSGDTLNKIAA 376

Qy        123 DTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKGKGENGSSTDE-GEGLVTAK 181
                  :  |: | :      : :| |  | |   |   || | |  ||: :: |
Db        377 QYGVSVANLRSWNGISGDLIFVGQKLIVKKGASGN--TG-GSGSGGSNNNQSGTNTYYTV 433

Qy        182 EVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTKVTF---ASGN--GTTATVSKDDQG 236
              :  | :||   :   : || ::      : :  | |:     ||||  |:    | ::|
Db        434 KSGDTLNKIAAQYGVSVANLRSWNGISGDLIFVGQKLIVKKGASGNTGGSNNGGSNNNQS 493

Qy        237 NITVKYDVNVGDALN-------VNQLQNSGWNLDSKAVAGSSGKVI-----SGNVSPSKG 284
                | :  || ||        |:      ||  |  :  :  |:|     |||   |
Db        494 GTNTYYTIKSGDTLNKIAAQYGVSVANLRSWNGISGDLIFAGQKIIVKKGTSGNTGGSS- 552

Qy        285 KMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSV--DDEGAL--- 339
                    | |:|    ::| |          : :|   |     |    :  |:|
Db        553 --------NGGSN--NNQSGTNTYYTIKSGDTLNKISAQFGVSVANLQAWNNISGSLIFA 602

Qy        340 ----------NVGSKDANKPVRITNVAPG------VKEGDVTNVAQLKGVAQNLNNRIDN 383
                       | || : |||   ||   |      :| ||  |
Db        603 GQKIIVKKGANSGSTNTNKP---TNNGGGATTSYTIKSGDTLN----------------K 643

Qy        384 VNGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYLGEAGYAIGYSSISAGGNWIIKGT 443
              ::      :|   :  |:       |  :   | | : : |        |||||     :
Db        644 ISAQFGVSVANLRSWNGI-------KGDLIFAGQTIIVKKG-------ASAGGNASSTNS 689

Qy        444 ASG 446
              |||
Db        690 ASG 692
```

```
RESULT 6
BIGA_SALTY
ID   BIGA_SALTY     STANDARD;      PRT;  1953 AA.
AC   P25927; P25928; Q9XCQ3;
DT   01-MAY-1992 (Rel. 22, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Putative surface-exposed virulence protein bigA precursor.
GN   BIGA OR STM3478.
OS   Salmonella typhimurium.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Salmonella.
OX   NCBI_TaxID=602;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 14028;
RA   Stojiljkovic I., Valentine P., Heffron F.;
RT   "Salmonella typhimurium rhs homolog.";
RL   Submitted (MAR-1999) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX   MEDLINE=21534948; PubMed=11677609;
RA   McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA   Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA   Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA   Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA   Waterston R., Wilson R.K.;
RT   "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT   LT2.";
RL   Nature 413:852-856(2001).
RN   [3]
RP   SEQUENCE OF 1-765 FROM N.A.
RC   STRAIN=LT2;
RX   MEDLINE=91100301; PubMed=1987123;
RA   Wu J.Y., Siegel L.M., Kredich N.M.;
RT   "High-level expression of Escherichia coli NADPH-sulfite reductase:
RT   requirement for a cloned cysG plasmid to overcome limiting siroheme
RT   cofactor.";
RL   J. Bacteriol. 173:325-333(1991).
CC   -!- CAUTION: Ref.3 sequence differs from that shown due to frameshifts
CC       in positions 414 and 732.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AF133696; AAD39458.1; -.
DR   EMBL; AE008859; AAL22340.1; -.
DR   EMBL; M64606; AAA27042.1; ALT_FRAME.
DR   EMBL; M64606; AAA27043.1; ALT_FRAME.
DR   StyGene; SG10437; bigA.
KW   Virulence; Repeat; Signal; Complete proteome.
FT   SIGNAL        1     27       POTENTIAL.
FT   CHAIN        28   1953       PUTATIVE SURFACE-EXPOSED VIRULENCE
FT                                PROTEIN BIGA.
FT   DOMAIN      101    252       15 X 11 AA TANDEM REPEATS.
FT   REPEAT      101    103       1 (INCOMPLETE).
FT   REPEAT      104    113       2 (INCOMPLETE).
FT   REPEAT      114    122       3 (INCOMPLETE).
FT   REPEAT      123    133       4.
FT   REPEAT      134    144       5.
FT   REPEAT      145    155       6.
FT   REPEAT      156    166       7.
FT   REPEAT      167    177       8.
FT   REPEAT      178    188       9.
FT   REPEAT      189    199       10.
```

```
FT   REPEAT      200    210       11.
FT   REPEAT      211    221       12.
FT   REPEAT      222    232       13.
FT   REPEAT      233    243       14.
FT   REPEAT      244    252       15 (INCOMPLETE).
FT   CONFLICT    207    207       D -> DRGDDDVTPPDD (IN REF. 1).
FT   CONFLICT    514    514       A -> R (IN REF. 3).
FT   CONFLICT   1698   1698       D -> N (IN REF. 1).
FT   CONFLICT   1795   1798       QYLE -> ITLQ (IN REF. 1).
FT   CONFLICT   1836   1837       SA -> T (IN REF. 1).
SQ   SEQUENCE   1953 AA;  200150 MW;  611B3F1C954D91AE CRC64;

  Query Match             7.6%;  Score 178;  DB 1;  Length 1953;
  Best Local Similarity   24.7%;  Pred. No. 0.088;
  Matches  131;  Conservative   65;  Mismatches  171;  Indels  164;  Gaps   33;

Qy         33 NFAKETAGTNGDT---------------TVHLNGIGSTLT-------------DMLLNTG 64
              ||| : | : | |                 |  ::| ||| |              || :  |
Db        429 NFAGDIAVSGGGTAIIIDGDNATIKNTGTSDISGAGSTGTVIDGNNARVNNDGDMTITDG 488

Qy         65 AT-TNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSAD 123
               |  ::| |||  |            |||         || |               ||
Db        489 GTGGHITGDNVVID------------NAG--------STTVS---------------GAD 513

Qy        124 TKTTTVNVESKDN------GKKTEVKIGAKTSVIKEKDGK-------LVTGKG------K 164
                 | : :|  ||      | :| :  ||  : |   |           | | |
Db        514 --ATALYIEG-DNALVINEGNQT-ISGGAVGTRIDGDDAHTTNTGDIAVDGAGSAAVIIN 569

Qy        165 GENGSSTDEGEGLVT--AKEVI--DAVNKAGWRMKTTTANGQTGQADKFETVTSGTKVTF 220
              |:||| |  |: |||  |  :|     |:|       | |     ||    | :| | ||
Db        570 GDNGSLTQAGDLLVTDGAMGIITYGTGNEA-----KNTGNATVRDADSVGFVVAGEKNTF 624

Qy        221 -------ASGNGTTATVSKDDQGNITVKYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGK 273
                    | ||| | ||  |   :|:  |:||   ::|  :   :    :|:| |
Db        625 KNKGDIDVSLNGTGALVS-GDMSQVTLDGDINV---VSVQDSEGVFSSATGVSVSGDSNA 680

Qy        274 V-ISGNVSPS---------KGKMDETVNINAGNNIEITRNGK-NI---DIATSMTPQFSS 319
              | |:|||: |         |    |  :  ||   :| ||  ||   |:: :
Db        681 VDITGNVNISADYGQDDLAAGAPPLTGVVVGGNGNTVTLNGALNIDDNDLSATGGQYLDV 740

Qy        320 VSLGAGADAPTLSVDDEGALNVGSKDANKPVRITNVAPGVKEGDVTNVAQLKG---VAQN 376
              | |    |   : :|  | :|:    :  |:  |:        |:| :: : |   |  |
Db        741 VGLSVTGDDNDVEID--GGINI--THSEDPLDGTS-------ADITGIS-VSGNSTVTLN 788

Qy        377 LNNRIDN---VNGN---ARAGIAQAI-----ATAGLVQAYLP------GKSMMAIGGGTY 419
               :: ||    | |:   ||     ::     :   :  :|:|           :|| | ||
Db        789 GHSTIDTNTVVGGHVVLARVNNGGSLILGDDSVVDVNVSYIPTGYYTYNALLMADGEGTS 848

Qy        420 LGEAG-------YAI----GYSSISAGGNWIIKGTASGNSRGHFGASASVG 459
              :   |       |::      | :|  |: ::  |:| :|      : : |
Db        849 IENKGDITSHGVYSVIRADNGSEVSNSGDILVYATSSNSSEDRAAITRASG 899


RESULT 7
AG43_ECOLI
ID   AG43_ECOLI     STANDARD;      PRT;  1039 AA.
AC   P39180; P75614; P76360; P97241; Q46771;
DT   01-FEB-1995 (Rel. 31, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Antigen 43 precursor (AG43) (Fluffing protein).
GN   FLU OR B2000.
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12 / MG1655;
RX   MEDLINE=97426617; PubMed=9278503;
RA   Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA   Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA   Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA   Mau B., Shao Y.;
RT   "The complete genome sequence of Escherichia coli K-12.";
RL   Science 277:1453-1474(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12;
RX   MEDLINE=97251358; PubMed=9097040;
RA   Itoh T., Aiba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K.,
RA   Kasai H., Kimura S., Kitakawa M., Kitagawa M., Makino K., Miki T.,
RA   Mizobuchi K., Mori H., Mori T., Motomura K., Nakade S., Nakamura Y.,
RA   Nashimoto H., Nishio Y., Oshima T., Saito N., Sampei G., Seki Y.,
RA   Sivasundaram S., Tagami H., Takeda J., Takemoto K., Wada C.,
RA   Yamamoto Y., Horiuchi T.;
RT   "A 460-kb DNA sequence of the Escherichia coli K-12 genome
RT   corresponding to the 40.1-50.0 min region on the linkage map.";
RL   DNA Res. 3:379-392(1996).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ML 308-225;
RA   Henderson I.R., Owen P.;
RL   Submitted (JAN-1997) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   PRELIMINARY SEQUENCE OF 53-78.
RC   STRAIN=ML 308-225;
RX   MEDLINE=89291704; PubMed=2661530;
RA   Caffrey P., Owen P.;
RT   "Purification and N-terminal sequence of the alpha subunit of antigen
RT   43, a unique protein complex associated with the outer membrane of
RT   Escherichia coli.";
RL   J. Bacteriol. 171:3634-3640(1989).
RN   [5]
RP   SEQUENCE OF 53-63.
RC   STRAIN=K12 / EMG2;
RX   MEDLINE=97443975; PubMed=9298646;
RA   Link A.J., Robison K., Church G.M.;
RT   "Comparing the predicted and observed properties of proteins encoded
RT   in the genome of Escherichia coli K-12.";
RL   Electrophoresis 18:1259-1313(1997).
RN   [6]
RP   GENE NAME.
RX   MEDLINE=97257509; PubMed=9103983;
RA   Henderson I.R., Meehan M., Owen P.;
RT   "Antigen 43, a phase-variable bipartite outer membrane protein,
RT   determines colony morphology and autoaggregation in Escherichia coli
RT   K-12.";
RL   FEMS Microbiol. Lett. 149:115-120(1997).
CC   -!- FUNCTION: CONTROLS COLONY FORM VARIATION AND AUTOAGGREGATION. MAY
CC       FUNCTION AS AN ADHESIN.
CC   -!- SUBUNIT: HETERODIMER OF THE PERIPHERAL MEMBRANE PROTEIN (ALPHA
CC       CHAIN) ANCHORED TO THE INTEGRAL OUTER MEMBRANE PROTEIN (BETA
CC       CHAIN).
CC   -!- SUBCELLULAR LOCATION: Outer membrane-associated.
CC   -!- SIMILARITY: TO ADHESIN AIDI-I AND TO BORDETELLA PERTACTIN.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE000291; AAC75061.1; ALT_INIT.
DR   EMBL; D90838; BAA15825.1; ALT_INIT.
DR   EMBL; D90839; BAA15832.1; ALT_INIT.
DR   EMBL; U24429; AAB47869.1; -.
DR   HSSP; P07505; 1SRD.
DR   EcoGene; EG12686; flu.
DR   InterPro; IPR006315; Autotransport.
DR   InterPro; IPR005546; Autotransporter.
DR   InterPro; IPR004899; Pertactin.
DR   Pfam; PF03797; Autotransporter; 1.
DR   Pfam; PF03212; Pertactin; 1.
```

```
DR   TIGRFAMs; TIGR01414; autotrans_barl; 1.
KW   Outer membrane; Signal; Complete proteome.
FT   SIGNAL        1     52
FT   CHAIN        53    551       ANTIGEN 43 ALPHA CHAIN.
FT   CHAIN       552   1039       ANTIGEN 43 BETA CHAIN.
FT   VARIANT       2      2       K -> N (IN STRAIN ML 308-225).
FT   VARIANT      41     42       SL -> FF (IN STRAIN ML 308-225).
FT   VARIANT      46     46       T -> K (IN STRAIN ML 308-225).
FT   VARIANT     157    157       W -> L (IN STRAIN ML 308-225).
FT   VARIANT     188    188       V -> F (IN STRAIN ML 308-225).
FT   VARIANT     303    305       ATN -> STI (IN STRAIN ML 308-225).
FT   VARIANT     320    320       A -> T (IN STRAIN ML 308-225).
FT   VARIANT     372    372       N -> Q (IN STRAIN ML 308-225).
FT   VARIANT     493    493       E -> V (IN STRAIN ML 308-225).
FT   VARIANT     497    497       S -> N (IN STRAIN ML 308-225).
FT   VARIANT     585    585       H -> Y (IN STRAIN ML 308-225).
FT   VARIANT     709    709       E -> K (IN STRAIN ML 308-225).
FT   VARIANT     721    721       M -> T (IN STRAIN ML 308-225).
FT   VARIANT     751    753       GHL -> SHF (IN STRAIN ML 308-225).
FT   VARIANT     803    803       S -> P (IN STRAIN ML 308-225).
FT   VARIANT     815    815       A -> V (IN STRAIN ML 308-225).
FT   VARIANT     824    824       C -> S (IN STRAIN ML 308-225).
FT   VARIANT     829    835       LNLVHTS -> MNLIYNA (IN STRAIN ML 308-
FT                                225).
FT   VARIANT     845    847       QGT -> LGA (IN STRAIN ML 308-225).
FT   VARIANT     855    855       S -> T (IN STRAIN ML 308-225).
FT   VARIANT     888    888       Q -> L (IN STRAIN ML 308-225).
FT   VARIANT    1025   1025       S -> I (IN STRAIN ML 308-225).
FT   CONFLICT     61     63       ETV -> TTT (IN REF. 5).
SQ   SEQUENCE   1039 AA;  106841 MW;  5170D647C8DEEBE0 CRC64;

  Query Match             7.5%;  Score 177;  DB 1;  Length 1039;
  Best Local Similarity   20.8%;  Pred. No. 0.048;
  Matches  114;  Conservative   84;  Mismatches  206;  Indels  144;  Gaps   28;

Qy          5 GLINVETEKLSFG-ANGKKVNIISDTKGLNFAKET-----------AGTNGDTTVHLNGI 52
              |:     :: ||  ||  :     : || :  :               ||  |||   |:
Db         66 GTLANHDNQIVFGTTNGMTI-----STGLEYGPDNEANTGGQWVQDGGTANKTTVTSGGL 120

Qy         53 -----GSTLTDMLLNTGATTNV----TNDNVTDDEK---KRAASVKDVLN-AGWNIKGVK 99
                   | :::| ::: |   ::      |  :   |:   : | :  |:|  || :  ||
Db        121 QRVNPGGSVSDTVISAGGGQSLQGRAVNTTLNGGEQWMHEGAIATGTVINDKGWQV--VK 178

Qy        100 PGTTASDNVDFVRT----------YDTVEFLSADTKTTTVN------VESKDNGKKTEVK 143
              ||| |:| |  | |           || :|:  |   ||:|      | ::     | |
Db        179 PGTVATDTV--VNTGAEGGPDAENGDTGQFVRGDAVRTTINKNGRQIVRAEGTANTTVVY 236

Qy        144 IGAKTSV-----------------------IKEKDGKLVTGKGKGENGSSTDEGEGLV 178
                |   :|                     :   ||  :   | |  |::|   :| :
Db        237 AGGDQTVHGHALDTTLNGGYQYVHNGGTASDTVVNSDGWQIVKNG-GVAGNTTVNQKGRL 295

Qy        179 TAKEVIDAVN---KAGWRMKTTTANGQT-----------GQADKFETVTSGTKVTFASG 223
                  | |   | |  : |:||   |            |:||    : :| ::   :|
Db        296 QVDAGGTATNVTLKQGGALVTSTAATVTGINRLGAFSVVEGKADNV-VLENGGRLDVLTG 354

Qy        224 NGTTATVSKDDQGNITVKYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSK 283
              :  | |   || | : |:   | | |  |: : | |       :   ||  :|| |  |
Db        355 HTATNT-RVDDGGTLDVR---NGGTATTVS-MGNGG------VLLADSGAAVSGTRSDGK 403

Qy        284 -----GKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGA 338
                    |   : :  |:: :       | ||  |       :   |  | | ::::
Db        404 AFSIGGGQADALMLEKGSSFTLNAG----DTATDTTVNGGLFTARGGTLAGTTTLNNGAI 459

Qy        339 LNVGSKDANKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVNGNARAGIAQAIAT 398
              | :  |  |      |    ::|||    | |:| :   |  ::  :|:    ::    |
Db        460 LTLSGKTVN------NDTLTIREGD----ALLQGGSLTGNGSVEK-SGSGTLTVSNTTLT 508

Qy        399 AGLVQAYLPGKSMMAIGGGT----YLGEAGYAIGYS-------------SISAGGNWII 440
                |      :  :  :   |     : : | |:  :               ::::|  | |
Db        509 QKAVNL---NEGTLTLNDSTVTTDVIAQRGTALKLTGSTVLNGAIDPTNVTLASGATWNI 565

Qy        441 KGTASGNS 448
                |:  |
```

```
Db        566 PDNATVQS 573


RESULT 8
OMPB_RICJA
ID   OMPB_RICJA     STANDARD;      PRT;  1656 AA.
AC   O06653;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Outer membrane protein B precursor (168 kDa surface-layer protein)
DE   (Surface protein antigen) (Cell surface antigen 5) (Sca5) (rOmpB)
DE   (rOmp B) [Contains: 120 kDa surface-exposed protein (Surface protein
DE   antigen) (120 kDa outer membrane protein ompB); 32 kDa beta peptide].
GN   OMPB.
OS   Rickettsia japonica.
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC   Rickettsiaceae; Rickettsieae; Rickettsia.
OX   NCBI_TaxID=35790;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=YH;
RA   Uchiyama T.;
RT   "Sequencing of the gene encoding the protein rOmp B of Rickettsia
RT   japonica.";
RL   Submitted (MAY-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: THE 120 KDA SURFACE-EXPOSED PROTEIN IS A MAJOR
CC       STRUCTURAL PROTEIN WHICH MAY PLAY A ROLE AS A RICKETTSIAL
CC       VIRULENCE FACTOR AND/OR IMMUNOGEN DURING INFECTION (BY
CC       SIMILARITY).
CC   -!- FUNCTION: THE 32 KDA BETA PEPTIDE MAY SERVE AS A MEMBRANE ANCHOR
CC       (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: CELL WALL. THIS BACTERIUM IS COVERED BY A S-
CC       LAYER WITH HEXAGONAL SYMMETRY.
CC   -!- SIMILARITY: BELONGS TO THE RICKETTSIAE OMPA/OMPB FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AB003681; BAA20138.1; -.
DR   InterPro; IPR006315; Autotransport.
DR   InterPro; IPR005546; Autotransporter.
DR   Pfam; PF03797; Autotransporter; 1.
DR   TIGRFAMs; TIGR01414; autotrans_barl; 2.
KW   Antigen; S-layer; Cell wall.
FT   CHAIN         1   1338       120 kDa SURFACE-EXPOSED PROTEIN.
FT   CHAIN      1339   1656       32 kDa BETA PEPTIDE.
FT   DOMAIN      528    533       POLY-GLY.
SQ   SEQUENCE   1656 AA;  168097 MW;  3132A69C9DD5999F CRC64;

  Query Match             7.5%;  Score 177;  DB 1;  Length 1656;
  Best Local Similarity   22.3%;  Pred. No. 0.082;
  Matches  138;  Conservative   61;  Mismatches  212;  Indels  208;  Gaps   30;

Qy          1 TDETGLINVETEKLSFGANG------------------KKVNIISDTKGLNFAKETAGT 41
              | :  ||  :   :  ||||                   | :| | | :|  |
Db        193 TQKAPLILADNALIVNGANGTLNVTNGFIQVSDKSFATVKAIN-IGDGQGFMFNTNATNA 251

Qy         42 N------GDTTVHLNGIGSTLTDMLLN-TGATTNVTNDNVTDDEKKRAASVKDVLNAGWN 94
              |      | ||:: ||   |   :||:  || ||:    |||             : | |
Db        252 NALNLQAGGTTINFNGTDGTGRLVLLSKNGAATDF---NVTG------------SLGGN 295

Qy         95 IKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVN--------VESKDNGKKTEV--KI 144
              :||:    | : |   :          |:    | |        | | ||||   :  ::
Db        296 LKGIIELNTVAINGQLIANAG-----PANAVIGTNNGAGRAAGFVVSVDNGKAATIDGQV 350

Qy        145 GAKTSVIKE----------------------KDGKLVTGKGKGENGSSTDEGEGLVTAK 181
                || ||:                       |   :   : |  :|| | |
```

```
Db        351 YAKDMVIQSANANGQVNFRHIVDVGIDGTTAFKTAASIVAITQNSNFGTTDFG-NLAAQV 409

Qy        182 EVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTKVTFASGNGTTATVSKDDQGNITVK 241

Db        410 TVPDTM--------TLTGN-FTGDANN-PGNTAGV-ITFAA-NGTLASASAD--ANVAVT 455

Qy        242 YDVNVGDALNVNQLQNSGWNLDSKAV--AGSSGKVISGNVSPSKGKMDETV--------- 290

Db        456 NNITAIEASGVGVVQLSGTHTAELRLGNAGSVFKLADGTV--INGKVNQTVLVGGVLAAG 513

Qy        291 ------------------------------------------------NINA-GNNI 298

Db        514 AITLDGSATITGDIGNGGGGAALQSITLANDATKTLTLGGANIISANGGTINFQANGGTI 573

Qy        299 EITRNGKNI----DIATSMTPQFSSVSLGAGADAPTLSVDDE----------GALNVGS 343

Db        574 KLTSTQNNIVVDCDLAIA-TDQTGVVDASSLTNAQTLTISGTIGIIGANNTTLGQFNIGS 632

Qy        344 KDA---NKPVRITNVAPGVKEGDV---------TNVAQLKGVAQNLNNRIDNVNGNARAG 391

Db        633 SKTTLNGGNVAINELVIG-NNGSVQFAHNTYLITRTTNAAGQGKIIFNPVVNNNTTLAAG 691

Qy        392 IAQAIATAGLVQAYLPGK-----SMMAIGGGTYLGEAGYA---------IGYSSISAGGN 437

Db        692 TNLGSAANPLAEINFGSKGARADTVLNVGEGVNL----YATNITTTDANVGSFVFNAGGK 747

Qy        438 WIIKGTASGNSRGHFGASA 456

Db        748 NIVSGTVGGQQGNKFNTVA 766
```

```
RESULT 9
HLYA_PROMI
ID   HLYA_PROMI     STANDARD;      PRT;  1577 AA.
AC   P16466;
DT   01-AUG-1990 (Rel. 15, Created)
DT   01-AUG-1990 (Rel. 15, Last sequence update)
DT   01-NOV-1990 (Rel. 16, Last annotation update)
DE   Hemolysin precursor.
GN   HPMA.
OS   Proteus mirabilis.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Proteus.
OX   NCBI_TaxID=584;
RN   [1]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 30-43.
RC   STRAIN=Isolate 477-12;
RX   MEDLINE=90170827; PubMed=2407716;
RA   Uphoff T.S., Welch R.A.;
RT   "Nucleotide sequencing of the Proteus mirabilis calcium-independent
RT   hemolysin genes (hpmA and hpmB) reveals sequence similarity with the
RT   Serratia marcescens hemolysin genes (shlA and shlB).";
RL   J. Bacteriol. 172:1206-1216(1990).
CC   -!- FUNCTION: BACTERIAL HEMOLYSINS ARE EXOTOXINS THAT ATTACK BLOOD
CC       CELL MEMBRANES AND CAUSE CELL RUPTURE BY MECHANISMS NOT CLEARLY
CC       DEFINED.
CC   -!- FUNCTION: CELL-BOUND HEMOLYSIN, WHICH RELEASES HEME-IRON FROM
CC       ERYTHROCYTES BY INTERACTION WITH THE ERYTHROCYTE MEMBRANE. HPMA
CC       REQUIRES HPMB FUNCTION.
CC   -!- SUBCELLULAR LOCATION: Outer membrane.
CC   -!- MISCELLANEOUS: THE CONSERVED AMPHIPATHIC DOMAINS IN SHLA AND HPMA
CC       MAY BE RESPONSIBLE FOR PORE FORMATION.
CC   -!- SIMILARITY: TO S.MARCESCENS HEMOLYSIN (SHLA).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M30186; AAA25657.1; -.
DR   PIR; A35140; A35140.
KW   Hemolysis; Toxin; Outer membrane; Signal.
FT   SIGNAL        1     29
FT   CHAIN        30   1577       HEMOLYSIN.
SQ   SEQUENCE   1577 AA;  165869 MW;  175975E0C924B2D9 CRC64;

  Query Match             7.4%;  Score 173.5;  DB 1;  Length 1577;
  Best Local Similarity   22.6%;  Pred. No. 0.12;
  Matches  132;  Conservative   85;  Mismatches  168;  Indels  199;  Gaps   35;

Qy          8 NVETEKLSFGANGKKVNIISDTKGLNFA----KETAGT--------------------NG 43

Db        711 STETEQANSTISGANVDLQAN-KDVTFAGSDLKTTAGNASITGDNVAFVSTENKKQTDNT 769

Qy         44 DTTVH--------LNGIGS----------TLTDMLLNTGATTNVTND-NVTDDEKKRAAS 84

Db        770 DTTISGGFSYTGGVDKVGSKADFQYDKQHTQTEVTKNRGSQTEVAGDLTITAN------- 822

Qy         85 VKDVLNAGWNIKGVKPGTTASDNVD--FVRTYDTVEFL----SADTKTTTVNV------- 131

Db        823 -KDLLHEG----------ASHHVEGRYQESGENIQHLAVNDSETSKTDSLNVGIDVGVN 870

Qy        132 ----------ESKDNGKKTEVKIGAKTSVIKE---KD----------------GKLVTG 161

Db        871 LDYSGVTKPVKKAIEDGVNT-TKPGNNTDLTKKVTARDAIANLANLSNLETPNVGVEVGI 929

Qy        162 KGKGENGSSTDE---GEGLVTAKEVIDAVNK---AGWRMKTT------TANGQTGQA--D 207

Db        930 KGGGSQQSQTDSQAVSTSINAGKIDIDSNNKLHDQGTHYQSTQEGISLTANTHTSEATLD 989

Qy        208 KFETVTSGTK----VTFASGNGTTATVSKDDQGNITVKYDVNVGDALNVNQLQNSGW--N 261

Db        990 KHQTTFHETKGGGQIGVSTKTGSDITVAIKGEGQTT-------DNALMETKAKGSQFTSN 1042

Qy        262 LDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITR------------NGK-NID 308

Db       1043 GDISINVGENAHYEGAQFDAQKGK---TV-INAGGDLTLAQATDTHSESQSNVNGSANLK 1098

Qy        309 IATSMTPQFSSVSLGAGADAPTL-SVDDEGALNVGSKDANKPVRITNVAPGVKEGDVTNV 367

Db       1099 VGT--TPE--SKDYGGGFNAGTTHHSKEQTTAKVGTITGSQGIELNAGHNLTLQG--THL 1152

Qy        368 AQLKGVAQNLNNRID------------NVNGNARAGIAQAIATAGLVQAYLPGKSM--- 411

Db       1153 SSEQDIALNATNKVDLQSASSEHTEKGNNLSGGVQAGF--------------GKKMTDD 1197

Qy        412 -MAIGGGTYLGEAGYAIG---YSSISAGGNWIIKG---TASGNS 448

Db       1198 ASSVNG---LGSAQFAIGKQDEKSVSREGGTINNSGNLTINGNS 1238
```

```
RESULT 10
ICEN_XANCT
ID   ICEN_XANCT     STANDARD;      PRT;  1567 AA.
AC   P18127;
DT   01-NOV-1990 (Rel. 16, Created)
DT   01-NOV-1990 (Rel. 16, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Ice nucleation protein.
GN   INAX.
OS   Xanthomonas campestris (pv. translucens).
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC   Xanthomonadaceae; Xanthomonas.
OX   NCBI_TaxID=343;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=X56S;
RX   MEDLINE=91080859; PubMed=2259339;
RA   Zhao J., Orser C.S.;
RT   "Conserved repetition in the ice nucleation gene inaX from
RT   Xanthomonas campestris pv. translucens.";
RL   Mol. Gen. Genet. 223:163-166(1990).
CC   -!- FUNCTION: ICE NUCLEATION PROTEINS ENABLE BACTERIA TO NUCLEATE
```

```
CC       CRYSTALLIZATION IN SUPERCOOLED WATER.
CC   -!- SUBCELLULAR LOCATION: Outer membrane (By similarity).
CC   -!- DOMAIN: CONTAINS 153 IMPERFECT REPEATS OF THE CONSENSUS
CC       OCTAPEPTIDE A-G-Y-G-S-T-L-T; FURTHER ON A 16-RESIDUE AND A
CC       REGIONAL 48-RESIDUE PERIODICITY IS SUPERIMPOSED.
CC   -!- MISCELLANEOUS: A STRUCTURAL MODEL IS SUGGESTED IN WHICH THE ICE
CC       NUCLEATION PROTEIN DISPLAYS A SYMMETRY RELATED TO THAT OF ICE.
CC   -!- SIMILARITY: BELONGS TO THE BACTERIAL ICE NUCLEATION PROTEIN
CC       FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X52970; CAA37140.1; -.
DR   HSSP; P06620; 1INA.
DR   InterPro; IPR000258; Ice_nucleatn.
DR   Pfam; PF00818; Ice_nucleation; 81.
DR   PRINTS; PR00327; ICENUCLEATN.
DR   PROSITE; PS00314; ICE_NUCLEATION; 57.
KW   Ice nucleation; Repeat; Outer membrane.
SQ   SEQUENCE   1567 AA;  152548 MW;  C8B451D959ECAD63 CRC64;

  Query Match             7.4%;  Score 173;  DB 1;  Length 1567;
  Best Local Similarity   20.8%;  Pred. No. 0.12;
  Matches   91;  Conservative   67;  Mismatches  220;  Indels   60;  Gaps   15;

Qy         38 TAGTNGDTTVHLNGIGSTLTDMLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKG 97
              : || |  :  ::| ||| |      |: :::|    :    :: : |     ||:  |
Db        864 STGTAGADSTLISGYGSTQT-----AGSDSSLTAGYGSTQTARKGSDV----TAGYGSTG 914

Qy         98 VKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGK 157
                 ||  :  :  | : :  :|:  |     ::   | :::  |  ::      |
Db        915 -----TAGADSTLIAGYGSTQTSGSDSSLTAGYGSTQTARKGSDMTAGYGSTGTAGADST 969

Qy        158 LVTGKGKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTK 217
              |: | |  :   |              |:  ||:    |   |    |   | |:|
Db        970 LIAGYGSTQTSGS-------------DSSLTAGYGSTQTAREGSDVTAGYGSTGTAGAD 1015

Qy        218 VTFASGNGTTATVSKDDQGNITVKYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISG 277
               |  :| |:| |   |   ::|  |    |      |  :      |   ||:   :|:|
Db       1016 STLIAGYGSTQTAGSD--SSLTAGY----GSTQTARQGSDVTAGYGSTGTAGADSTLIAG 1069

Qy        278 NVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEG 337
               |      | ::    |:  :  | | :|      | : |    ||||:  ::    |
Db       1070 YGSTQTAGSDSSLTAGYGST-QTARQGSDI------TAGYGSTGT-AGADSSLIA--GYG 1119

Qy        338 ALNVGSKDANKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVNGNARAGIAQAIA 397
              :      |:|      :     ::  :|       | : :: |         ||    :
Db       1120 STQTAGYDSNLTAGYGSTQTAREDSSLTAGYGSTSTAGHDSSLIAGYGSTQTAGYNSILT 1179

Qy        398 TA-GLVQAYLPGKSMMAIGGGTYLGEAGY----AIGYSSISAGG--NWIIKGTASGNSRG 450
              |  |  |      |: |  | |    |||      || |    |  : :  |  | :: |
Db       1180 TGYGSTQTAQESSSLTAGYGST--STAGYDSTLTAGYGSTQTAGYKSTLTAGYGSNSTAG 1237

Qy        451 H-------FGASASVGYQ 461
              |       :|::   ||:
Db       1238 HESSLIAGYGSTQIAGYE 1255


RESULT 11
SLAP_CAMFE
ID   SLAP_CAMFE     STANDARD;      PRT;   933 AA.
AC   P35827;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   01-OCT-1994 (Rel. 30, Last annotation update)
DE   S-layer protein (Surface array protein) (SAP).
GN   SAPA.
OS   Campylobacter fetus.
OC   Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
OC   Campylobacteraceae; Campylobacter.
OX   NCBI_TaxID=196;
RN   [1]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 1-20.
RC   STRAIN=84-32 / 23D;
RX   MEDLINE=90354448; PubMed=2387868;
RA   Blaser M.J., Gotschlich E.C.;
RT   "Surface array protein of Campylobacter fetus. Cloning and gene
RT   structure.";
RL   J. Biol. Chem. 265:14529-14535(1990).
RN   [2]
RP   ERRATUM.
RX   MEDLINE=91035477; PubMed=2229082;
RA   Blaser M.J., Gotschlich E.C.;
RL   J. Biol. Chem. 265:19372-19372(1990).
CC   -!- FUNCTION: THE S-LAYER IS A PARACRYSTALLINE MONO-LAYERED ASSEMBLY
CC       OF PROTEINS WHICH COAT THE SURFACE OF BACTERIA. THIS PROTEIN IS
CC       CRITICAL FOR VIRULENCE.
CC   -!- SUBCELLULAR LOCATION: CELL WALL. THIS BACTERIUM IS COVERED BY A
CC       S-LAYER WITH HEXAGONAL SYMMETRY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; J05577; AAA23032.1; -.
KW   Cell wall; S-layer.
SQ   SEQUENCE   933 AA;  96757 MW;  F88C729B4BA5B1E9 CRC64;

  Query Match             7.3%;  Score 172.5;  DB 1;  Length 933;
  Best Local Similarity   21.8%;  Pred. No. 0.072;
  Matches  127;  Conservative   59;  Mismatches  219;  Indels  177;  Gaps   24;

Qy          1 TDETGLI------NVET-----------------EKLSFGANGKKVNIISD-TKGLNFA 35
              ||:  ||      |||                 ||: |   |  |:  :| :| :|
Db        260 TDQAELIITKRRTNVENINIISDLETSGDFVFNGYEKVGFNVLGDIVSFATDASKSVNV- 318

Qy         36 KETAGT----------------------NGDTTVHLNGIGSTLTDMLLNTGATTNVTN 71
               || ||                     :  |:|:|     |:|    |   : |:
Db        319 -ETTGTITAFTAAGTGKVDVVAGKISALTADSRTSVNLTATNDTITLTSANAATSVNLKQ 377

Qy         72 DNVTD-----------DEKKRAA----------SVKDVLNAGWNIKGVKPGTTASDNV 108
                |           : : | |          :||   |   |   |  |   ||
Db        378 RQAKDATITSAMQQKYNNRRNRIATITSATAVENLTVKHATNVALNGGMDKLATVTLDNA 437

Qy        109 DFVRTYDTVEFLSADTKTTTVN-VESKDNG-------------KKTEVKIGAKTSVIKE 153
                  |           :|:| : |  ||                |:   |:   |:   :
Db        438 ALTAAIDI-------KSASTLNLINSSVNGPKHLYSSKRRYCKFKRAAAKVKLNTTAATD 490

Qy        154 K---------DGKLVTGKGKGENGSSTDEGEG---LVTAKEVIDAVNKAGWRMKTTTAN- 200
              :         |  |       :  | |  | |   ::   ||   ||     : ||  |
Db        491 QTVTLKANATDNSLEFDSATAKTTSVTASGSGKTLVIKGAEVETLVN-----IDTTAFNA 545

Qy        201 ------GQTGQADKF--ETVTSGTKVTFASGNGTTATVSK--DDQGNITVKYDVNVGDAL 250
                  |:|||  ||  :| |   |: |   ||| |      |  || |:          :
Db        546 LQSVSFGKTGQGGKFSVKTGTGDDKIEFV---GTTLTEGSVIDAPGNDTIAMKSAALTSA 602

Qy        251 NVNQLQNSGWNLDSKAVAG---SSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNI 307
              |   ::|      | |||    ||    :  :  :|   | |: ||       |
Db        603 NFTMIKNIENVAISDAVATADLSSSAFKNSVIITTKEAADTTLTINKDQVI-------NF 655

Qy        308 DIATSMTPQFSSVSLGAGADAPTLSVDDEGALNVGSKDANKPVRITNVAPGVKEGDVTNV 367
                | : : :  :| |    |   | :  :  |:  :||       ||:| |    |   |
Db        656 TAADAGSVKLITVKLN---DVTALMIVVKIVLDAAAKD-------TNIALGTAAADKALV 705

Qy        368 AQLKGVAQNLNNRIDNVNGNARAGIAQAIATAGLVQAYLPGKSMMAIGG---------- 416
                    ::  |: :|  :          ||  | | |   : : | |
```

```
Db        706 ---------IDTGIETLNITSLVKATSPETTANTVNAKLTDVTSIIIDGMQITLGHAGTA 756

Qy        417 GTYLG----------EAGYAIGYSSISAGGNWIIKGTASGNS 448
              ||             :||     |:|: | |  ||| :  :|
Db        757 GTDYSKVSMIDASALKAGLTFDASAITLGANATIKGGSGADS 798


RESULT 12
120K_RICRI
ID   120K_RICRI     STANDARD;      PRT;  1300 AA.
AC   P14914;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   120 kDa surface-exposed protein.
GN   P120.
OS   Rickettsia rickettsii.
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC   Rickettsiaceae; Rickettsieae; Rickettsia.
OX   NCBI_TaxID=783;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=R;
RX   MEDLINE=90136087; PubMed=2515418;
RA   Gilmore R.D. Jr., Joste N., McDonald G.A.;
RT   "Cloning, expression and sequence analysis of the gene encoding the
RT   120 kD surface-exposed protein of Rickettsia rickettsii.";
RL   Mol. Microbiol. 3:1579-1586(1989).
CC   -!- FUNCTION: MAJOR STRUCTURAL PROTEIN WHICH MAY PLAY A ROLE AS
CC       RICKETTSIAL VIRULENCE FACTOR AND/OR IMMUNOGEN DURING INFECTION.
CC   -!- SUBCELLULAR LOCATION: CELL WALL. THIS BACTERIUM IS COVERED BY A
CC       S-LAYER WITH HEXAGONAL SYMMETRY.
CC   -!- MISCELLANEOUS: ITS C-TERMINUS POTENTIALLY MAY BEAR THE EPITOPES
CC       CONFERRING ANTIGENICITY TO THE PROTEIN.
CC   -!- SIMILARITY: BELONGS TO THE RICKETTSIAE OMPA/OMPB FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X16353; CAA34402.1; -.
DR   PIR; S07575; S07575.
DR   InterPro; IPR006315; Autotransport.
DR   InterPro; IPR005546; Autotransporter.
DR   Pfam; PF03797; Autotransporter; 1.
DR   TIGRFAMs; TIGR01414; autotrans_barl; 2.
KW   Antigen; Glycoprotein; Cell wall; S-layer.
FT   CARBOHYD      7      7       N-LINKED (POTENTIAL).
FT   CARBOHYD     66     66       N-LINKED (POTENTIAL).
FT   CARBOHYD     86     86       N-LINKED (POTENTIAL).
FT   CARBOHYD    103    103       N-LINKED (POTENTIAL).
FT   CARBOHYD    147    147       N-LINKED (POTENTIAL).
FT   CARBOHYD    268    268       N-LINKED (POTENTIAL).
FT   CARBOHYD    330    330       N-LINKED (POTENTIAL).
FT   CARBOHYD    375    375       N-LINKED (POTENTIAL).
FT   CARBOHYD    415    415       N-LINKED (POTENTIAL).
FT   CARBOHYD    424    424       N-LINKED (POTENTIAL).
FT   CARBOHYD    430    430       N-LINKED (POTENTIAL).
FT   CARBOHYD    436    436       N-LINKED (POTENTIAL).
FT   CARBOHYD    444    444       N-LINKED (POTENTIAL).
FT   CARBOHYD    515    515       N-LINKED (POTENTIAL).
FT   CARBOHYD    547    547       N-LINKED (POTENTIAL).
FT   CARBOHYD    593    593       N-LINKED (POTENTIAL).
FT   CARBOHYD    655    655       N-LINKED (POTENTIAL).
FT   CARBOHYD    698    698       N-LINKED (POTENTIAL).
FT   CARBOHYD    710    710       N-LINKED (POTENTIAL).
FT   CARBOHYD    799    799       N-LINKED (POTENTIAL).
FT   CARBOHYD    800    800       N-LINKED (POTENTIAL).
FT   CARBOHYD    826    826       N-LINKED (POTENTIAL).
FT   CARBOHYD    844    844       N-LINKED (POTENTIAL).
FT   CARBOHYD    861    861       N-LINKED (POTENTIAL).
FT   CARBOHYD    879    879       N-LINKED (POTENTIAL).
FT   CARBOHYD    920    920       N-LINKED (POTENTIAL).
FT   CARBOHYD    926    926       N-LINKED (POTENTIAL).
FT   CARBOHYD   1116   1116       N-LINKED (POTENTIAL).
FT   CARBOHYD   1128   1128       N-LINKED (POTENTIAL).
FT   CARBOHYD   1140   1140       N-LINKED (POTENTIAL).
FT   CARBOHYD   1146   1146       N-LINKED (POTENTIAL).
FT   CARBOHYD   1211   1211       N-LINKED (POTENTIAL).
SQ   SEQUENCE   1300 AA;  132801 MW;  E09E52C3F647243D CRC64;

  Query Match             7.2%;  Score 170;  DB 1;  Length 1300;
  Best Local Similarity   24.0%;  Pred. No. 0.14;
  Matches  119;  Conservative   63;  Mismatches  190;  Indels  124;  Gaps   29;

Qy         37 ETAGTNGDTTV-HLNGIGSTLTDMLLNTGATTNVTNDN--VTDDEKKRAASVK--DVLNA 91
              ::|   |     |:  :|:  |       :   :| |:     |    || :|  : :
Db          4 QSANATGQVNFRHIVDVGADGTTAFKTAASKVTITQDSNFGNTDFGNLAAQIKVPNAITL 63

Qy         92 GWNIKG--VKPGTTASDNVDFVRTYD---TVEFLSADTKTTTVN----VESKDNG----- 137
              |  |    || ||      | |:|   |:|  |||      |    :|:   |
Db         64 TGNFTGDASNPGNTAG-----VITFDANGTLESASADANVAVTNNITAIEASGAGVVQLS 118

Qy        138 --KKTEVKIGAKTSVIKEKDGKLVTGK--------GKGENGSSTDEGEGLVTAKEVIDAV 187
                 |:::|   |: |  || :: ||         |    |: | :|   :|    |
Db        119 GTHAAELRLGNAGSIFKLADGTVINGKVNQTALVGGALAAGTITLDGSATITGD--IGNA 176

Qy        188 NKAGWRMKTTTANGQTGQADKFETVTSGTKVTFASGNGT-----------TATVSKDDQ 235
                |    : | ||      |  :|:| |      :| ||            |:|     |
Db        177 GGAAALQRITLAN------DAKKTLTLGGANIIGAGGGTIDLQANGGTIKLTST-----Q 225

Qy        236 GNITVKYDV-------NVGDALNVNQLQNSGWN-------LDSKAVA----GSSGKVIS- 276
               || | :|:        | || ::   |    |        ::| :    |||  |:|
Db        226 NNIVVDFDLAIATDQTGVVDASSLTNAQTLTINGKIGTIGANNKTLGQFNIGSSKTVLSN 285

Qy        277 GNVSPSKGKMDETVNINAG------NNIEITR------NGKNI--DIATSMTPQFSSVSL 322
               |||:     ::| |  | |      :   |||      || |   :  : |   :  :|
Db        286 GNVA-----INELVIGNDGAVQFAHDTYLITRTTNAAGQGKIIFNPVVNNGTTLAAGTNL 340

Qy        323 GAGADAPTLSVDDEGALNVGSKDANKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRID 382
              |:       : :    :| |||  |    : ||  ||     ||:        |: : :
Db        341 GS-------ATNPLAEINFGSKGVNVDT-VLNVGEGVNL-YATNITTTDA---NVGSFVF 388

Qy        383 NVNG-NARAGIAQAIATAGLVQ-------AYLPGKSMMAIGGGTYLGEAGYAIGYSSISA 434
              | | |  :|      | |  |        |   | ::  :|  |: |    |   |::
Db        389 NAGGTNIVSG------TVGGQQGNKFNTVALENGTTVKFLGNATFNGNTTIAAN-STLQI 441

Qy        435 GGNWIIKGTASGNSRG 450
              |||:     || :  |
Db        442 GGNYTADCVASADGTG 457


RESULT 13
OMPA_RICCN
ID   OMPA_RICCN     STANDARD;      PRT;  2021 AA.
AC   Q52657; P95591; P95592; P95593; P95594; Q52667; Q52668; Q52669;
AC   Q52670; Q52674;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Outer membrane protein A precursor (190 kDa antigen) (Cell surface
DE   antigen) (rOmpA) (rOmp A).
GN   OMPA OR RC1273.
OS   Rickettsia conorii.
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC   Rickettsiaceae; Rickettsieae; Rickettsia.
OX   NCBI_TaxID=781;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Malish 7;
RX   MEDLINE=94171067; PubMed=8125327;
RA   Crocquet-Valdes P.A., Weiss K., Walker D.H.;
```

```
RT   "Sequence analysis of the 190-kDa antigen-encoding gene of Rickettsia
RT   conorii (Malish 7 strain).";
RL   Gene 140:115-119(1994).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Malish 7;
RX   MEDLINE=21442074; PubMed=11557893;
RA   Ogata H., Audic S., Renesto-Audiffren P., Fournier P.-E., Barbe V.,
RA   Samson D., Roux V., Cossart P., Weissenbach J., Claverie J.-M.,
RA   Raoult D.;
RT   "Mechanisms of evolution in Rickettsia conorii and R. prowazekii.";
RL   Science 293:2093-2098(2001).
RN   [3]
RP   SEQUENCE OF 8-204 FROM N.A.
RC   STRAIN=Indian tick typhus, M1, Malish 7, and Moroccan;
RX   MEDLINE=97015921; PubMed=8862558;
RA   Roux V., Fournier P.E., Raoult D.;
RT   "Differentiation of spotted fever group rickettsiae by sequencing and
RT   analysis of restriction fragment length polymorphism of PCR-amplified
RT   DNA of the gene encoding the protein rOmpA.";
RL   J. Clin. Microbiol. 34:2058-2065(1996).
RN   [4]
RP   SEQUENCE OF 953-2012 FROM N.A.
RC   STRAIN=Indian tick typhus, M1, Malish 7, and Moroccan;
RA   Raoult D., Fournier P.E., Roux V.;
RT   "Phylogenetic analysis of spotted fever group rickettsiae by study
RT   of the outer surface protein rOmpA.";
RL   Submitted (DEC-1996) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: ELICITS PROTECTIVE IMMUNITY (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: CELL WALL. THIS BACTERIUM IS COVERED BY A
CC       S-LAYER WITH HEXAGONAL SYMMETRY.
CC   -!- PTM: GLYCOSYLATED (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE RICKETTSIAE OMPA/OMPB FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; U01028; AAA17405.1; -.
DR   EMBL; AE008674; AAL03811.1; -.
DR   EMBL; U43794; AAB49549.1; -.
DR   EMBL; U43798; AAB49550.1; -.
DR   EMBL; U43806; AAB49551.1; -.
DR   EMBL; U45244; AAB49566.1; -.
DR   EMBL; U46918; AAA86663.1; -.
DR   EMBL; U83440; AAC35176.1; -.
DR   EMBL; U83443; AAC35179.1; -.
DR   EMBL; U83448; AAC35184.1; -.
DR   EMBL; U83453; AAC35189.1; -.
DR   InterPro; IPR006315; Autotransport.
DR   Pfam; PF03797; Autotransporter; 1.
DR   TIGRFAMs; TIGR01414; autotrans_barl; 1.
KW   Antigen; Repeat; Signal; Cell wall; S-layer; Glycoprotein;
KW   Complete proteome.
FT   SIGNAL       1     38       POTENTIAL.
FT   CHAIN       39   2021       OUTER MEMBRANE PROTEIN A.
FT   DOMAIN     238    946       THR-RICH.
FT   DOMAIN    1424   1528       THR-RICH.
FT   VARIANT     60     60       N -> NN (IN STRAIN INDIAN TICK TYPHUS).
FT   VARIANT     76     76       R -> H (IN STRAIN INDIAN TICK TYPHUS).
FT   VARIANT     86    137       MISSING (IN STRAIN M1).
FT   VARIANT    126    133       MISSING (IN STRAIN MOROCCAN).
FT   VARIANT    953    954       VT -> II (IN STRAIN INDIAN TICK TYPHUS).
FT   VARIANT   1245   1245       D -> A (IN STRAINS INDIAN TICK TYPHUS, M1
FT                               AND MOROCCAN).
FT   VARIANT   1308   1308       N -> H (IN STRAIN MOROCCAN).
FT   VARIANT   1877   1877       M -> I (IN STRAIN INDIAN TICK TYPHUS).
FT   CONFLICT    10     10       Q -> K (IN REF. 1).
FT   CONFLICT    92     92       I -> V (IN REF. 1).
FT   CONFLICT   126    126       V -> I (IN REF. 1).
FT   CONFLICT   137    137       T -> N (IN REF. 1).
FT   CONFLICT   157    157       G -> D (IN REF. 1).
FT   CONFLICT   368    369       IS -> VN (IN REF. 1).
FT   CONFLICT   374    388       KATLGGAIIKATTTK -> LLQVQGGVVKANTIN (IN
FT                               REF. 1).
FT   CONFLICT   640    640       N -> D (IN REF. 1).
FT   CONFLICT   669    669       V -> I (IN REF. 1).
FT   CONFLICT   793    793       N -> D (IN REF. 1).
FT   CONFLICT   803    804       VN -> IS (IN REF. 1).
FT   CONFLICT   809    823       LLRVQGGVVKSNTIN -> KATLGGAIIKATTTK (IN
FT                               REF. 1).
FT   CONFLICT   898    898       D -> Y (IN REF. 1).
FT   CONFLICT   908    908       P -> N (IN REF. 1).
FT   CONFLICT   985    985       N -> K (IN REF. 1).
FT   CONFLICT  1009   1009       L -> S (IN REF. 1).
FT   CONFLICT  1013   1013       Y -> S (IN REF. 1).
FT   CONFLICT  1182   1182       K -> Q (IN REF. 1).
FT   CONFLICT  1314   1314       N -> Y (IN REF. 4).
FT   CONFLICT  1451   1451       H -> N (IN REF. 1).
FT   CONFLICT  1624   1624       G -> D (IN REF. 1).
FT   CONFLICT  1628   1628       E -> G (IN REF. 1).
FT   CONFLICT  1872   1872       A -> V (IN REF. 1).
FT   CONFLICT  1875   1875       T -> P (IN REF. 1).
FT   CONFLICT  1878   1879       MS -> LP (IN REF. 1).
FT   CONFLICT  1936   1936       E -> A (IN REF. 1).
FT   CONFLICT  1965   1970       MTAPLP -> ITPPLS (IN REF. 1).
FT   CONFLICT  1997   1997       G -> R (IN REF. 1).
SQ   SEQUENCE   2021 AA;  203328 MW;  327FC42D7CB24668 CRC64;

  Query Match             7.1%;  Score 168;  DB 1;  Length 2021;
  Best Local Similarity   21.8%;  Pred. No. 0.29;
  Matches  134;  Conservative   65;  Mismatches  209;  Indels  208;  Gaps   30;

Qy          2 DETGLINVETEKLSFGANGKKVNIISDTKGLNFAKETAGTNGDTTVHLNGIGSTLTDMLL 61
              | ||  |          ||    |:: |           ||   ||:: | |    :  :
Db        625 DNTGNAN----------NG----IVTFTGNSTVTGNIGNTNALATVNV-GAGIATLEGAV 669

Qy         62 NTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLS 121
                || :||          |||| : |    : |    ||  |||  :    :  ::
Db        670 IKATTTKLTN----------AASVLTLTNVNAVLTGAIDNTTGVDNVGVLNLNGALSQVT 719

Qy        122 ADTKTT----TVNV---ESKDNG---KKTEVKIGAKTSVIKEKDGKLVTG----KGKGEN 167
               :   |    |::|   ::   |   | |  |:    | :   :  :|||     |   |
Db        720 GNIGNTNALATISVGAGKATLGGAVIKATTTKLTDNASAVTFTNPVVVTGAIDNTGNANN 779

Qy        168 GSSTDEGEGLVTAK----EVIDAVNKA-------GWRMKTTTAN---------------- 200
              | :|  |:  ||        :  ||         |  :|: | |
Db        780 GIATFTGDSTVTGNIGNTNALATVNVGAGLLRVQGGVVKSNTINLTDNASAVTFTNPVVV 839

Qy        201 ----GQTGQADK-FETVTSGTKVTFASGN-GTTATVS--------------------KDD 234
                  || |:    | |  : ||   ||    ||:|                      |:
Db        840 TGAIDNTGNANNGIVTFTGDSTVTGNIGNTNALATISVGAGKATLGGAIIKATTTKLTDN 899

Qy        235 QGNITVKYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINA 294
                :|    | |  |::     |:| | ::  |  :    ::||:  :      |||: |
Db        900 ASAVTFTNPVVVTGAID-----NTG-NANNGIVTFTGDSTVTGNIGNTNAL--ATVNVGA 951

Qy        295 GNNIEI--TRNGKNIDIATSMTPQFSSVSLGAGADAP-----TLSVDDEGALNVGSK--- 344
              |  ::   : :  |||     | :|:    | |   |       ::  :   ||| :|
Db        952 GVTLQAGGSLDANNIDFGARSTLEFNGPLDGGGNAIPYYFKGAIANGNNAILNVNTKLLT 1011

Qy        345 ----------------------DANK-PVRITN--------------------------- 354
                                   ||:   | | |
Db       1012 AYHLTIGTVAEINIGAGNLFAIDASAGDVTILNAQDIHFRALDSALVLSNLTGVGVNNIL 1071

Qy        355 -----VAPGVKEGDV------------TNVAQLKGVAQNLNNRIDNVNGNA--------R 389
                  ||||| || |             :|||   | |:|    | :| ||
Db       1072 LAADLVAPGVDEGTVVFDGGVNGLNIGSNVA---GAARN----IGDVGGNKFNTLLIYNA 1124

Qy        390 AGIAQAIATAGLVQAYLPGK----SMMAIGGGT-YLGEAGYAI----GYSSISAGG---- 436
              |   :   |:    :        |  |   ||  : :| | |    |  :| ||
Db       1125 VTITDDVNLEGIQNVLINNNADFTSSTAFNAGTIQINDATYTIDANNGNLNIPAGNIKFA 1184
```

```
Qy        437 ---NWIIKGTASGNSR 449
                  :|   :||| |
Db       1185 HADAQLILQNSSGNDR 1200


RESULT 14
OMPB_RICTY
ID   OMPB_RICTY     STANDARD;      PRT;  1645 AA.
AC   P96989;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Outer membrane protein B precursor (168 kDa surface-layer protein)
DE   (Surface protein antigen) (Cell surface antigen 5) (Sca5) (rOmpB)
DE   (rOmp B) [Contains: 120 kDa surface-exposed protein (Surface protein
DE   antigen) (120 kDa outer membrane protein ompB); 32 kDa beta peptide].
GN   OMPB OR SLP.
OS   Rickettsia typhi.
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC   Rickettsiaceae; Rickettsieae; Rickettsia.
OX   NCBI_TaxID=785;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Wilmington;
RX   MEDLINE=94040787; PubMed=8224886;
RA   Hahn M.-J., Kim K.-K., Kim I., Chang W.-H.;
RT   "Cloning and sequence analysis of the gene encoding the crystalline
RT   surface layer protein of Rickettsia typhi.";
RL   Gene 133:129-133(1993).
RN   [2]
RP   PARTIAL SEQUENCE.
RC   STRAIN=Wilmington;
RX   MEDLINE=92114896; PubMed=1370573;
RA   Ching W.M., Carl M., Dasch G.A.;
RT   "Mapping of monoclonal antibody binding sites on CNBr fragments of
RT   the S-layer protein antigens of Rickettsia typhi and Rickettsia
RT   prowazekii.";
RL   Mol. Immunol. 29:95-105(1992).
RN   [3]
RP   IDENTIFICATION OF CLEAVAGE SITE.
RX   MEDLINE=92104668; PubMed=1729180;
RA   Hackstadt T., Messer R., Cieplak W., Peacock M.G.;
RT   "Evidence for proteolytic cleavage of the 120-kilodalton outer
RT   membrane protein of rickettsiae: identification of an avirulent
RT   mutant deficient in processing.";
RL   Infect. Immun. 60:159-165(1992).
CC   -!- FUNCTION: THE 120 kDa SURFACE-EXPOSED PROTEIN IS A MAJOR
CC       STRUCTURAL PROTEIN WHICH MAY PLAY A ROLE AS A RICKETTSIAL
CC       VIRULENCE FACTOR AND/OR IMMUNOGEN DURING INFECTION.
CC   -!- FUNCTION: THE 32 kDa BETA PEPTIDE MAY SERVE AS A MEMBRANE ANCHOR.
CC   -!- SUBCELLULAR LOCATION: CELL WALL. THIS BACTERIUM IS COVERED BY A
CC       S-LAYER WITH HEXAGONAL SYMMETRY.
CC   -!- SIMILARITY: BELONGS TO THE RICKETTSIAE OMPA/OMPB FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L04661; AAB48987.1; -.
DR   PIR; JN0896; JN0896.
DR   InterPro; IPR006315; Autotransport.
DR   InterPro; IPR005546; Autotransporter.
DR   Pfam; PF03797; Autotransporter; 1.
DR   TIGRFAMs; TIGR01414; autotrans_barl; 1.
KW   Antigen; S-layer; Transmembrane; Cell wall.
FT   CHAIN         1   1353       120 kDa SURFACE-EXPOSED PROTEIN.
FT   CHAIN      1354   1645       32 kDa BETA PEPTIDE.
FT   TRANSMEM   1415   1429       MEMBRANE ANCHOR (POTENTIAL).
FT   CONFLICT    657    657       H -> N (IN REF. 2).
FT   CONFLICT    842    842       V -> I (IN REF. 2).
FT   CONFLICT   1071   1071       G -> A (IN REF. 2).
FT   CONFLICT   1306   1306       G -> S (IN REF. 2).
SQ   SEQUENCE   1645 AA;  169698 MW;  0CB5641C7EB185EE CRC64;

  Query Match             7.1%;  Score 166;  DB 1;  Length 1645;
  Best Local Similarity   22.4%;  Pred. No. 0.29;
  Matches  136;  Conservative   59;  Mismatches  209;  Indels  204;  Gaps   31;

Qy          1 TDETGLINVETEKLSFGANG------------------KKVNIISDTKGLNFAKETAGT 41
              | |  |   :  |:  ||||                  | :| | | :|| |
Db        193 TQEAPLTLGDNAKIVNGANGILNITNGFVKVSDKTFAGIKTIN-IGDNQGLMFNTTPDAA 251

Qy         42 N------GDTTVHLNGIGSTLTDMLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNI 95
              |      |  |:: ||     |  |:      | |  |||              : | |:
Db        252 NALNLQGGGNTINFNGRDG--TGKLVLVSKNGNATEFNVTG------------SLGGNL 296

Qy         96 KGV-KPGTTA----------------SDN-----VDFVRTYDTVEFLSADTKTTTVNVE 132
              ||| :  |||                :||      |: : |    : :  | :  |
Db        297 KGVIEFDTTAAAGKLIANGGAANAVIGTDNGAGRAAGFIVSVD-----NGNAATISGQVY 351

Qy        133 SKD--------NGKKT-----EVKIGAKTSVIKEKDGKLVTGKGKGENGS--STDEG--- 174
              :||          |: |     :| :| ||:  |  | |::      || |  ||| |
Db        352 AKDIVIQSANAGGQVTFEHLVDVGLGGKTN-FKTADSKVII----TENASFGSTDFGNLA 406

Qy        175 -------EGLVTAKEVIDAVNK---AGWRMKTTTANGQ--TGQADKFETVTS-------G 215
                    ::|   : || |    ||  : |  |||   :|  |    ||:       |
Db        407 VQIVVPNNKILTGNFIGDAKNNGNTAG--VITFNANGTLVSGNTDPNIVVTNIKAIEVEG 464

Qy        216 TKVTFASG-------NGTTATVSKDDQGNITVKYDVNVGDALNVNQLQNSGWNLDSKAVA 268
              :   ||        |   :: |   | : :   ||    :| | |          ||  |:
Db        465 AGIVQLSGIHGAELRLGNAGSIFKLADGTV-INGPVNQNPLVNNNALAAGSIQLDGSAII 523

Qy        269 GSSGKVISGNVSPSKGKMDET----------------VNINAGNNIEITRNGKNI----- 307
              :| :  || : :    | |                 :  |||  |    ||  |
Db        524 --TGDI--GNGAVNAALQDITLANDASKILTLSGANIIGANAGGAIHFQANGGTIQLTST 579

Qy        308 ----------DIATSMTPQFSSVSL------------GAGADAPTLSVDDEGALNVGSK 344
                       |: |  |    : ||               ||:  ||     |  ||||
Db        580 QNNILVDFDLDVTTDQTGVVDASSLTNNQTLTINGSIGTIGANTKTL-----GRFNVGSS 634

Qy        345 ----------------DANKPVRIT-NVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVNGN 387
                           : :  | :| |     |  :  |  ::   |  :|      :|
Db        635 KTILNAGDVAINELVMENDGSVHLTHNTYLITKTINAANQGKIIVAADPINTDTALADGT 694

Qy        388 ARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYLGEAGYA---------IGYSSISAGGNW 438
                    ::          | |:: || |  |    ||         :|     :||
Db        695 NLGSAESPLSNIHFATKAANGDSILHIGKGVNL----YANNITTTDANVGSLHFRSGGTS 750

Qy        439 IIKGTASG 446
              |: ||  |
Db        751 IVSGTVGG 758


RESULT 15
YDEK_ECOLI
ID   YDEK_ECOLI     STANDARD;      PRT;  1325 AA.
AC   P32051; P76140; P77168;
DT   01-OCT-1993 (Rel. 27, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Hypothetical lipoprotein ydeK precursor (ORFT).
GN   YDEK OR ORFT OR B1510.
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12 / MG1655;
RX   MEDLINE=97426617; PubMed=9278503;
RA   Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
```

```
RA   Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA   Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA   Mau B., Shao Y.;
RT   "The complete genome sequence of Escherichia coli K-12.";
RL   Science 277:1453-1474(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12;
RX   MEDLINE=97251357; PubMed=9097039;
RA   Aiba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K., Itoh T.,
RA   Kasai H., Kashimoto K., Kimura S., Kitakawa M., Kitagawa M.,
RA   Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,
RA   Nakade S., Nakamura Y., Nashimoto H., Nishio Y., Oshima T., Saito N.,
RA   Sampei G., Seki Y., Sivasundaram S., Tagami H., Takeda J.,
RA   Takemoto K., Takeuchi Y., Wada C., Yamamoto Y., Horiuchi T.;
RT   "A 570-kb DNA sequence of the Escherichia coli K-12 genome
RT   corresponding to the 28.0-40.1 min region on the linkage map.";
RL   DNA Res. 3:363-377(1996).
RN   [3]
RP   SEQUENCE OF 595-1325 FROM N.A.
RX   MEDLINE=94100243; PubMed=8274505;
RA   Cartwright P.J., Timms M.W., Lithgow T., Hoej P.B., Hoogenraad N.J.;
RT   "An Escherichia coli gene showing a potential ancestral relationship
RT   to the genes for the mitochondrial import site proteins ISP42 and
RT   MOM38.";
RL   Biochim. Biophys. Acta 1153:345-347(1993).
CC   -!- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor
CC       (Potential).
CC   -!- SIMILARITY: TO E.COLI YFAL.
CC   -!- SIMILARITY: SOME, TO FUNGAL MITOCHONDRIAL IMPORT SITE PROTEINS
CC       ISP42 AND MOM38.
CC   -!- CAUTION: REF.3 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO A
CC       FRAMESHIFT IN POSITION 653.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE000248; AAC74583.1; -.
DR   EMBL; D90793; BAA15190.1; ALT_INIT.
DR   EMBL; D90794; BAA15197.1; ALT_INIT.
DR   EMBL; X73295; CAA51730.1; ALT_FRAME.
DR   PIR; A64905; A64905.
DR   EcoGene; EG11780; ydeK.
DR   PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
KW   Hypothetical protein; Membrane; Lipoprotein; Signal;
KW   Complete proteome.
FT   SIGNAL        1     18       POTENTIAL.
FT   CHAIN        19   1325       HYPOTHETICAL LIPOPROTEIN YDEK.
FT   LIPID        19     19       N-ACYL DIGLYCERIDE (POTENTIAL).
FT   CONFLICT    884    884       N -> K (IN REF. 3).
FT   CONFLICT   1317   1317       M -> S (IN REF. 3).
SQ   SEQUENCE   1325 AA;  136514 MW;  26A3A066FA19AD7D CRC64;

  Query Match             7.0%;  Score 165;  DB 1;  Length 1325;
  Best Local Similarity   24.3%;  Pred. No. 0.26;
  Matches  125;  Conservative   54;  Mismatches  218;  Indels  117;  Gaps   27;

Qy          4 TGLINVETEKLSFGANGKKVNIISDTKGLNFAKETAGTNGDTTVHLNGIGST--LTD--- 58
              ||::|:       | |     | :   | |  ::: ||       :| :| |  | |
Db         90 TGVLNILA-----GGNASLTTITTSVIGAN--EDSEGT-------VNVLGGTWRLYDSGN 135

Qy         59 --MLLNTG----ATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVR 112
                || |     | |:      |           |  | :  ||  ||   :  | |
Db        136 NARPLNVGQSGTGTLNIKQKGHVDG---------GYLRLGSSTGGV--GTVNVEGEDSVL 184

Qy        113 TYDTVEFLSADTKTTTVNVESKDNGKKTE-------VKIGAKTSVIKEKDGK-LVTGK-- 162
              | :  |  | |  | :  :     | | |         : |:   |: || |: |:
Db        185 TTELFEIGSYGTGSLNIT----DKGYVTSSIVAILGYQAGSNGQVVVEKGGEWLIKNNDS 240

Qy        163 ------GKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGT 216
                   |    | :|    |||||:  |   |  |      | | |    |:   :|
Db        241 SIEFQIGNQGTGEATIREGGLVTAENTIIGGNATG----IGTLNVQ----DQDSVITVRR 292

Qy        217 KVTFASGNGTTATVSKDDQGNITVK-YD-VNVGDALN--VNQLQNSGWNL---------- 262
                  |||   ||:  : | |  | |  | | |  :  ||      ||
Db        293 LYNGYFGNG---TVNISNNGLINNKEYSLVGVQDGSHGVVNVTDKGHWNFLGTGEAFRYI 349

Qy        263 ------DSKAVAGSSGKVISGNVSPSKGKMDETVNINAGN------NIEITRNGKNIDIA 310
                   | :    | ||| || ::     | ||     ||      |  ||  | |:
Db        350 YIGDAGDGELNVSSEGKVDSGIITAG---MKET---GTGNITVKDKNSVITNLGTNLGYD 403

Qy        311 TSMTPQFSSVSLGAGADAPTLSVDDEGALNV-----GSKDANKPVRITNVAPGVKEGDVT 365
                  |:  |       :|   : |  ||     |  : || |  |    ||   :::
Db        404 GHGEMNISNQGLVVSNGGSSLGYGETGVGNVSITTGGMWEVNKNVYTTIGVAGVGNLNIS 463

Qy        366 NVAQLKGVAQNLNNRIDNVNGNARAGIAQAIATAGLVQAYLP--GKSMMAIGGGTYLGEA 423
              :    | |:||:     |  :|      :  | ::   |   :   |  :: :  |  |
Db        464 DGG--KFVSQNITFLGDKASGIGTLNLMDATSSFDTVGINVGNFGSGIVNVSNGATLNST 521

Qy        424 GYAI--------GYSSISAGGNWIIKGTASGNSR 449
              ||          |  :||    | :| |:| |::
Db        522 GYGFIGGNASGKGIVNISTDSLWNLK-TSSTNAQ 554


Search completed: October  6, 2003, 09:24:17
Job time : 10.1345 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: October 6, 2003, 09:13:10 ; Search time 33.9322 Seconds
(without alignments)
3513.485 Million cell updates/sec

Title: US-09-771-382-36
Perfect score: 2350
Sequence: 1 TDETGLINVETEKLSFGANG.....TASGNSRGHFGASASVGYQW 462

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_23:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2337	99.4	592	2	Q93QY2	Q93qy2 neisseria m
2	2299	97.8	592	2	Q9JPS9	Q9jps9 neisseria m
3	2280	97.0	594	2	Q9JPI3	Q9jpi3 neisseria m
4	2280	97.0	594	2	Q9JPS2	Q9jps2 neisseria m
5	2261	96.2	598	2	Q9JPS0	Q9jps0 neisseria m
6	2261	96.2	598	2	Q93QY5	Q93qy5 neisseria m
7	2261	96.2	598	2	Q9JPT0	Q9jpt0 neisseria m
8	2260	96.2	590	2	Q9JPS3	Q9jps3 neisseria m
9	2260	96.2	599	2	Q9JPR8	Q9jpr8 neisseria m
10	2247	95.6	594	2	Q9JPH7	Q9jph7 neisseria m
11	2247	95.6	598	2	Q9JPR9	Q9jpr9 neisseria m
12	2228.5	94.8	600	2	Q9JPS6	Q9jps6 neisseria m
13	2224.5	94.7	526	2	Q9JPS4	Q9jps4 neisseria m
14	2224.5	94.7	530	2	Q9JPS1	Q9jps1 neisseria m
15	2213	94.2	594	2	Q93QY4	Q93qy4 neisseria m
16	2210.5	94.1	591	2	Q93QY3	Q93qy3 neisseria m
17	2210.5	94.1	591	2	Q9JPS7	Q9jps7 neisseria m
18	2210.5	94.1	591	16	Q9JR18	Q9jr18 neisseria m
19	2206.5	93.9	592	2	Q9AQF0	Q9aqf0 neisseria m
20	2156	91.7	600	2	Q9JPS5	Q9jps5 neisseria m
21	2136	90.9	592	16	Q9JQW4	Q9jqw4 neisseria m
22	2119	90.2	598	2	Q9JPR7	Q9jpr7 neisseria m
23	2116	90.0	589	2	Q93QY1	Q93qy1 neisseria m
24	2114	90.0	599	2	Q9JPS8	Q9jps8 neisseria m
25	2105	89.6	589	2	Q9JPI0	Q9jpi0 neisseria m
26	2104	89.5	595	2	Q9JPH0	Q9jph0 neisseria m
27	876	37.3	2353	2	P71401	P71401 haemophilus
28	851	36.2	1096	2	Q8GM79	Q8gm79 haemophilus
29	849.5	36.1	1098	2	Q48152	Q48152 haemophilus
30	832	35.4	1210	2	Q8GM74	Q8gm74 haemophilus
31	829	35.3	1210	2	Q8GM75	Q8gm75 haemophilus
32	821.5	35.0	1204	2	Q8GM76	Q8gm76 haemophilus
33	574.5	24.4	1002	2	Q8GM78	Q8gm78 haemophilus
34	573.5	24.4	1004	2	Q8GM77	Q8gm77 haemophilus
35	390.5	16.6	1299	16	Q9F3X6	Q9f3x6 pasteurella
36	372	15.8	2314	2	Q8KQM8	Q8kqm8 moraxella c
37	360	15.3	1588	16	Q8XDG4	Q8xdg4 escherichia
38	359.5	15.3	2059	16	Q9PD50	Q9pd50 xylella fas
39	357	15.2	1190	16	Q9PC04	Q9pc04 xylella fas
40	352.5	15.0	1461	16	Q8ZL64	Q8zl64 salmonella
41	351	14.9	1964	2	Q8KQM9	Q8kqm9 moraxella c
42	349	14.9	1778	16	Q8FCB2	Q8fcb2 escherichia
43	348.5	14.8	1107	16	Q9F2D8	Q9f2d8 salmonella
44	317.5	13.5	2712	16	Q9F3X5	Q9f3x5 pasteurella
45	317	13.5	641	16	Q8CKM1	Q8ckm1 yersinia pe

ALIGNMENTS

RESULT 1
Q93QY2
ID Q93QY2 PRELIMINARY; PRT; 592 AA.
AC Q93QY2;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE NhhA outer membrane protein.
GN NHHA.
OS Neisseria meningitidis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=487;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H41;
RA Peak I.R., Srikhanta Y., Dieckelman M., Moxon R., Jennings M.P.;
RT "Identification and characterization of a gene encoding a novel outer
RT membrane protein of Neisseria meningitidis.";
RL Submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF157609; AAK68870.1; -.
DR InterPro; IPR005594; YadA.
DR Pfam; PF03895; YadA; 1.
SQ SEQUENCE 592 AA; 61869 MW; F9403A0B4A18EEA7 CRC64;

Query Match 99.4%; Score 2337; DB 2; Length 592;
Best Local Similarity 99.8%; Pred. No. 2.7e-94;
Matches 460; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          2 DETGLINVETEKLSFGANGKKVNIISDTKGLNFAKETAGTNGDTTVHLNGIGSTLTDMLL 61
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        132 DLTGLINVETEKLSFGANGKKVNIISDTKGLNFAKETAGTNGDTTVHLNGIGSTLTDMLL 191

Qy         62 NTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLS 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        192 NTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLS 251

Qy        122 ADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKGKGENGSSTDEGEGLVTAK 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        252 ADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKGKGENGSSTDEGEGLVTAK 311

Qy        182 EVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTKVTFASGNGTTATVSKDDQGNITVK 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        312 EVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTKVTFASGNGTTATVSKDDQGNITVK 371

Qy        242 YDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEIT 301
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        372 YDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEIT 431

Qy        302 RNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGALNVGSKDANKPVRITNVAPGVKE 361
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        432 RNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGALNVGSKDANKPVRITNVAPGVKE 491

Qy        362 GDVTNVAQLKGVAQNLNNRIDNVNGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYLG 421
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        492 GDVTNVAQLKGVAQNLNNRIDNVNGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYLG 551

Qy        422 EAGYAIGYSSISAGGNWIIKGTASGNSRGHFGASASVGYQW 462
              |||||||||||||||||||||||||||||||||||||||||
Db        552 EAGYAIGYSSISAGGNWIIKGTASGNSRGHFGASASVGYQW 592


RESULT 2
Q9JPS9
ID   Q9JPS9      PRELIMINARY;      PRT;   592 AA.
AC   Q9JPS9;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Outer membrane protein GNA992.
GN   GNA992.
OS   Neisseria meningitidis.
OC   Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC   Neisseriaceae; Neisseria.
OX   NCBI_TaxID=487;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=860800;
RX   MEDLINE=20175756; PubMed=10710308;
RA   Pizza M., Scarlato V., Masignani V., Giuliani M.M., Arico' B.,
RA   Comanducci M., Jennings G.T., Baldi L., Bartolini E., Capecchi B.,
RA   Galeotti C.L., Luzzi E., Manetti R., Marchetti E., Mora M., Nuti S.,
RA   Ratti G., Santini L., Savino S., Scarselli M., Storni E., Zuo P.,
RA   Broeker M., Hundt E., Knapp B., Blair E., Mason T., Tettelin H.,
RA   Hood D.W., Jeffries A.C., Saunders N.J., Granoff D.M., Venter C.,
RA   Moxon E.R., Grandi G., Rappuoli R.;
RT   "Identification of Vaccine Candidates Against Serogroup B
RT   Meningococcus by Whole-Genome Sequencing.";
RL   Science 287:1816-1820(2000).
DR   EMBL; AF226361; AAF42510.1; -.
DR   InterPro; IPR005594; YadA.
DR   Pfam; PF03895; YadA; 1.
SQ   SEQUENCE   592 AA;  61917 MW;  4A3471514FD3C879 CRC64;

  Query Match            97.8%;  Score 2299;  DB 2;  Length 592;
  Best Local Similarity  98.3%;  Pred. No. 1.2e-92;
  Matches  453;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          2 DETGLINVETEKLSFGANGKKVNIISDTKGLNFAKETAGTNGDTTVHLNGIGSTLTDMLL 61
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        132 DLTGLINVETEKLSFGANGKKVNIISDTKGLNFAKETAGTNGDTTVHLNGIGSTLTDMLL 191

Qy         62 NTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLS 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        192 NTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLS 251

Qy        122 ADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKGKGENGSSTDEGEGLVTAK 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        252 ADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKGKGENGSSTDEGEGLVTAK 311

Qy        182 EVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTKVTFASGNGTTATVSKDDQGNITVK 241
              |||||||||||||||||||||||||||||||||| |||||| ||||||||||||||||
```

```
Db        312 EVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVM 371

Qy        242 YDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEIT 301
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        372 YDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEIT 431

Qy        302 RNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGALNVGSKDANKPVRITNVAPGVKE 361
              ||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||
Db        432 RNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDKGALNVGSKDANKPVRITNVAPGVKE 491

Qy        362 GDVTNVAQLKGVAQNLNNRIDNVNGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYLG 421
              ||||||||||||||||||||||:|||||||||||||||||||||||||||||||||| |
Db        492 GDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRG 551

Qy        422 EAGYAIGYSSISAGGNWIIKGTASGNSRGHFGASASVGYQW 462
              |||||||||||| ||||||||||||||||||||||||||||
Db        552 EAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 592


RESULT 3
Q9JPI3
ID   Q9JPI3      PRELIMINARY;      PRT;   594 AA.
AC   Q9JPI3;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Outer membrane protein GNA992.
GN   GNA992.
OS   Neisseria meningitidis.
OC   Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC   Neisseriaceae; Neisseria.
OX   NCBI_TaxID=487;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NG3/88, and BZ232;
RX   MEDLINE=20175756; PubMed=10710308;
RA   Pizza M., Scarlato V., Masignani V., Giuliani M.M., Arico' B.,
RA   Comanducci M., Jennings G.T., Baldi L., Bartolini E., Capecchi B.,
RA   Galeotti C.L., Luzzi E., Manetti R., Marchetti E., Mora M., Nuti S.,
RA   Ratti G., Santini L., Savino S., Scarselli M., Storni E., Zuo P.,
RA   Broeker M., Hundt E., Knapp B., Blair E., Mason T., Tettelin H.,
RA   Hood D.W., Jeffries A.C., Saunders N.J., Granoff D.M., Venter C.,
RA   Moxon E.R., Grandi G., Rappuoli R.;
RT   "Identification of Vaccine Candidates Against Serogroup B
RT   Meningococcus by Whole-Genome Sequencing.";
RL   Science 287:1816-1820(2000).
DR   EMBL; AF226376; AAF42525.1; -.
DR   EMBL; AF226369; AAF42518.1; -.
DR   InterPro; IPR005594; YadA.
DR   Pfam; PF03895; YadA; 1.
SQ   SEQUENCE   594 AA;  62086 MW;  1B25E03B90D04B46 CRC64;

  Query Match            97.0%;  Score 2280;  DB 2;  Length 594;
  Best Local Similarity  97.6%;  Pred. No. 8.1e-92;
  Matches  450;  Conservative    2;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          2 DETGLINVETEKLSFGANGKKVNIISDTKGLNFAKETAGTNGDTTVHLNGIGSTLTDMLL 61
              | | | :|||||||||||||||| ||||||||||||||||||||||||||||||| ||
Db        134 DLTDLTSVETEKLSFGANGKKVNITSDTKGLNFAKETAGTNGDTTVHLNGIGSTLTDTLL 193

Qy         62 NTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLS 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        194 NTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLS 253

Qy        122 ADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAK 181
              |||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db        254 ADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAK 313

Qy        182 EVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTKVTFASGNGTTATVSKDDQGNITVK 241
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        314 EVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTKVTFASGNGTTATVSKDDQGNITVK 373

Qy        242 YDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEIT 301
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        374 YDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEIT 433

Qy        302 RNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGALNVGSKDANKPVRITNVAPGVKE 361
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        434 RNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGALNVGSKDANKPVRITNVAPGVKE 493

Qy        362 GDVTNVAQLKGVAQNLNNRIDNVNGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYLG 421
                |||||||||||||||||| ||||:||||||||||||||||||||||||||||||||| |
Db        494 GDVTNVAQLKGVAQNLNNHIDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRG 553

Qy        422 EAGYAIGYSSISAGGNWIIKGTASGNSRGHFGASASVGYQW 462
                |||||||||||| ||||||||||||||||||||||||||
Db        554 EAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 594


RESULT 4
Q9JPS2
ID   Q9JPS2      PRELIMINARY;      PRT;   594 AA.
AC   Q9JPS2;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Outer membrane protein GNA992.
GN   GNA992.
OS   Neisseria meningitidis.
OC   Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC   Neisseriaceae; Neisseria.
OX   NCBI_TaxID=487;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NGE31;
RX   MEDLINE=20175756; PubMed=10710308;
RA   Pizza M., Scarlato V., Masignani V., Giuliani M.M., Arico' B.,
RA   Comanducci M., Jennings G.T., Baldi L., Bartolini E., Capecchi B.,
RA   Galeotti C.L., Luzzi E., Manetti R., Marchetti E., Mora M., Nuti S.,
RA   Ratti G., Santini L., Savino S., Scarselli M., Storni E., Zuo P.,
RA   Broeker M., Hundt E., Knapp B., Blair E., Mason T., Tettelin H.,
RA   Hood D.W., Jeffries A.C., Saunders N.J., Granoff D.M., Venter C.,
RA   Moxon E.R., Grandi G., Rappuoli R.;
RT   "Identification of Vaccine Candidates Against Serogroup B
RT   Meningococcus by Whole-Genome Sequencing.";
RL   Science 287:1816-1820(2000).
DR   EMBL; AF226379; AAF42528.1; -.
DR   InterPro; IPR005594; YadA.
DR   Pfam; PF03895; YadA; 1.
SQ   SEQUENCE   594 AA;  62114 MW;  1E2A63A78F53D256 CRC64;

  Query Match             97.0%;  Score 2280;  DB 2;  Length 594;
  Best Local Similarity   97.6%;  Pred. No. 8.1e-92;
  Matches  450;  Conservative    2;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          2 DETGLINVETEKLSFGANGKKVNIISDTKGLNFAKETAGTNGDTTVHLNGIGSTLTDMLL 61
              | | | :||||||||||||||||| ||||||||||||||||||||||||||||||| ||
Db        134 DLTDLTSVETEKLSFGANGKKVNITSDTKGLNFAKETAGTNGDTTVHLNGIGSTLTDTLL 193

Qy         62 NTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLS 121
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        194 NTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLS 253

Qy        122 ADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKGKGENGSSTDEGEGLVTAK 181
                |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        254 ADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAK 313

Qy        182 EVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTKVTFASGNGTTATVSKDDQGNITVK 241
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        314 EVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTKVTFASGNGTTATVSKDDQGNITVK 373

Qy        242 YDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEIT 301
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        374 YDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEIT 433

Qy        302 RNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGALNVGSKDANKPVRITNVAPGVKE 361
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        434 RNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGALNVGSKDANKPVRITNVAPGVKE 493

Qy        362 GDVTNVAQLKGVAQNLNNRIDNVNGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYLG 421
                |||||||||||||||||| ||||:||||||||||||||||||||||||||||||||| |
Db        494 GDVTNVAQLKGVAQNLNNHIDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRG 553

Qy        422 EAGYAIGYSSISAGGNWIIKGTASGNSRGHFGASASVGYQW 462
                |||||||||||| ||||||||||||||||||||||||||
Db        554 EAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 594


RESULT 5
Q9JPS0
ID   Q9JPS0      PRELIMINARY;      PRT;   598 AA.
AC   Q9JPS0;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Outer membrane protein GNA992 (NhhA outer membrane protein).
GN   GNA992 OR NHHA.
OS   Neisseria meningitidis.
OC   Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC   Neisseriaceae; Neisseria.
OX   NCBI_TaxID=487;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NGH15;
RX   MEDLINE=20175756; PubMed=10710308;
RA   Pizza M., Scarlato V., Masignani V., Giuliani M.M., Arico' B.,
RA   Comanducci M., Jennings G.T., Baldi L., Bartolini E., Capecchi B.,
RA   Galeotti C.L., Luzzi E., Manetti R., Marchetti E., Mora M., Nuti S.,
RA   Ratti G., Santini L., Savino S., Scarselli M., Storni E., Zuo P.,
RA   Broeker M., Hundt E., Knapp B., Blair E., Mason T., Tettelin H.,
RA   Hood D.W., Jeffries A.C., Saunders N.J., Granoff D.M., Venter C.,
RA   Moxon E.R., Grandi G., Rappuoli R.;
RT   "Identification of Vaccine Candidates Against Serogroup B
RT   Meningococcus by Whole-Genome Sequencing.";
RL   Science 287:1816-1820(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=H15;
RA   Peak I.R., Srikhanta Y., Dieckelman M., Moxon R., Jennings M.P.;
RT   "Identification and characterization of a gene encoding a novel outer
RT   membrane protein of Neisseria meningitidis.";
RL   Submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF226381; AAF42530.1; -.
DR   EMBL; AF157607; AAK68868.1; -.
DR   InterPro; IPR005594; YadA.
DR   Pfam; PF03895; YadA; 1.
SQ   SEQUENCE   598 AA;  62763 MW;  E6C7AEF0BB8A63CB CRC64;

  Query Match             96.2%;  Score 2261;  DB 2;  Length 598;
  Best Local Similarity   96.7%;  Pred. No. 5.5e-91;
  Matches  446;  Conservative    3;  Mismatches   12;  Indels    0;  Gaps    0;

Qy          2 DETGLINVETEKLSFGANGKKVNIISDTKGLNFAKETAGTNGDTTVHLNGIGSTLTDMLL 61
              | | | :||||||||||||| |||| |||||||||||||||| |||||||||||| ||
Db        138 DLTDLTSVETEKLSFGANGNKVNITSDTKGLNFAKETAGTNGDPTVHLNGIGSTLTDTLL 197

Qy         62 NTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLS 121
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        198 NTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLS 257

Qy        122 ADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKGKGENGSSTDEGEGLVTAK 181
                |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        258 ADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKGKDENGSSTDEGEGLVTAK 317

Qy        182 EVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTKVTFASGNGTTATVSKDDQGNITVK 241
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        318 EVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTKVTFASGNGTTATVSKDDQGNITVK 377

Qy        242 YDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEIT 301
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        378 YDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEIT 437

Qy        302 RNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGALNVGSKDANKPVRITNVAPGVKE 361
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        438 RNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGALNVGSKDANKPVRITNVAPGVKE 497

Qy        362 GDVTNVAQLKGVAQNLNNRIDNVNGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYLG 421
                 |||||||||||||||||||||:||||||||||||||||| ||||||||||||||||| |
Db        498 GDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLAQAYLPGKSMMAIGGGTYRG 557

Qy        422 EAGYAIGYSSISAGGNWIIKGTASGNSRGHFGASASVGYQW 462
                 ||||||||||||  |||:||||||||||||||||||||
Db        558 EAGYAIGYSSISDTGNWVIKGTASGNSRGHFGASASVGYQW 598


RESULT 6
Q93QY5
ID   Q93QY5      PRELIMINARY;      PRT;   598 AA.
AC   Q93QY5;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   NhhA outer membrane protein.
GN   NHHA.
OS   Neisseria meningitidis.
OC   Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC   Neisseriaceae; Neisseria.
OX   NCBI_TaxID=487;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BZ10;
RA   Peak I.R., Srikhanta Y., Dieckelman M., Moxon R., Jennings M.P.;
RT   "Identification and characterization of a gene encoding a novel outer
RT   membrane protein of Neisseria meningitidis.";
RL   Submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF157603; AAK68864.1; -.
DR   InterPro; IPR005594; YadA.
DR   Pfam; PF03895; YadA; 1.
SQ   SEQUENCE   598 AA;  62687 MW;  18CEFFE6410A15DF CRC64;

  Query Match             96.2%;  Score 2261;  DB 2;  Length 598;
  Best Local Similarity   96.5%;  Pred. No. 5.5e-91;
  Matches  445;  Conservative    4;  Mismatches   12;  Indels    0;  Gaps    0;

Qy          2 DETGLINVETEKLSFGANGKKVNIISDTKGLNFAKETAGTNGDTTVHLNGIGSTLTDMLL 61
                | | | :|||||||||||| |||| |||||||||||||||| ||||||||||||  ||
Db        138 DLTDLTSVETEKLSFGANGNKVNITSDTKGLNFAKETAGTNGDPTVHLNGIGSTLTDTLL 197

Qy         62 NTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLS 121
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        198 NTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLS 257

Qy        122 ADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKGKGENGSSTDEGEGLVTAK 181
                 |||||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||||
Db        258 ADTKTTTVNVESKDNGKRTEVKIGAKTSVIKEKDGKLVTGKGKGENGSSTDEGEGLVTAK 317

Qy        182 EVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTKVTFASGNGTTATVSKDDQGNITVK 241
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        318 EVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTKVTFASGNGTTATVSKDDQGNITVK 377

Qy        242 YDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEIT 301
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        378 YDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEIT 437

Qy        302 RNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGALNVGSKDANKPVRITNVAPGVKE 361
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        438 RNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGALNVGSKDANKPVRITNVAPGVKE 497

Qy        362 GDVTNVAQLKGVAQNLNNRIDNVNGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYLG 421
                 |||||||||||||||||||||:||||||||||||||||| ||||||||||||||||| |
Db        498 GDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLAQAYLPGKSMMAIGGGTYRG 557

Qy        422 EAGYAIGYSSISAGGNWIIKGTASGNSRGHFGASASVGYQW 462
                 ||||||||||||  |||:||||||||||||||||||||
Db        558 EAGYAIGYSSISDTGNWVIKGTASGNSRGHFGTSASVGYQW 598


RESULT 7
Q9JPT0
ID   Q9JPT0      PRELIMINARY;      PRT;   598 AA.
AC   Q9JPT0;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Outer membrane protein GNA992.
GN   GNA992.
OS   Neisseria meningitidis.
OC   Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC   Neisseriaceae; Neisseria.
OX   NCBI_TaxID=487;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=2996;
RX   MEDLINE=20175756; PubMed=10710308;
RA   Pizza M., Scarlato V., Masignani V., Giuliani M.M., Arico' B.,
RA   Comanducci M., Jennings G.T., Baldi L., Bartolini E., Capecchi B.,
RA   Galeotti C.L., Luzzi E., Manetti R., Marchetti E., Mora M., Nuti S.,
RA   Ratti G., Santini L., Savino S., Scarselli M., Storni E., Zuo P.,
RA   Broeker M., Hundt E., Knapp B., Blair E., Mason T., Tettelin H.,
RA   Hood D.W., Jeffries A.C., Saunders N.J., Granoff D.M., Venter C.,
RA   Moxon E.R., Grandi G., Rappuoli R.;
RT   "Identification of Vaccine Candidates Against Serogroup B
RT   Meningococcus by Whole-Genome Sequencing.";
RL   Science 287:1816-1820(2000).
DR   EMBL; AF226359; AAF42508.1; -.
DR   InterPro; IPR005594; YadA.
DR   Pfam; PF03895; YadA; 1.
SQ   SEQUENCE   598 AA;  62763 MW;  63A6A3BD7F0F2EE3 CRC64;

  Query Match             96.2%;  Score 2261;  DB 2;  Length 598;
  Best Local Similarity   96.5%;  Pred. No. 5.5e-91;
  Matches  445;  Conservative    4;  Mismatches   12;  Indels    0;  Gaps    0;

Qy          2 DETGLINVETEKLSFGANGKKVNIISDTKGLNFAKETAGTNGDTTVHLNGIGSTLTDMLL 61
                | | | :|||||||||||| |||| |||||||||||||||| ||||||||||||  ||
Db        138 DLTDLTSVETEKLSFGANGNKVNITSDTKGLNFAKETAGTNGDPTVHLNGIGSTLTDTLL 197

Qy         62 NTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLS 121
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        198 NTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLS 257

Qy        122 ADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKGKGENGSSTDEGEGLVTAK 181
                 |||||||||||||||:||||||||||||||||||||||||||||||||||||||||||||||||
Db        258 ADTKTTTVNVESKDNGKRTEVKIGAKTSVIKEKDGKLVTGKGKGENGSSTDEGEGLVTAK 317

Qy        182 EVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTKVTFASGNGTTATVSKDDQGNITVK 241
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        318 EVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTKVTFASGNGTTATVSKDDQGNITVK 377

Qy        242 YDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEIT 301
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        378 YDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEIT 437

Qy        302 RNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGALNVGSKDANKPVRITNVAPGVKE 361
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        438 RNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGALNVGSKDANKPVRITNVAPGVKE 497

Qy        362 GDVTNVAQLKGVAQNLNNRIDNVNGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYLG 421
                 |||||||||||||||||||||:||||||||||||||||| ||||||||||||||||| |
Db        498 GDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLAQAYLPGKSMMAIGGGTYRG 557

Qy        422 EAGYAIGYSSISAGGNWIIKGTASGNSRGHFGASASVGYQW 462
                 ||||||||||||  |||:||||||||||||||||||||
Db        558 EAGYAIGYSSISDTGNWVIKGTASGNSRGHFGTSASVGYQW 598
```

```
RESULT 8
Q9JPS3
ID   Q9JPS3      PRELIMINARY;      PRT;   590 AA.
AC   Q9JPS3;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Outer membrane protein GNA992.
GN   GNA992.
OS   Neisseria meningitidis.
OC   Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC   Neisseriaceae; Neisseria.
OX   NCBI_TaxID=487;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NGE28;
RX   MEDLINE=20175756; PubMed=10710308;
RA   Pizza M., Scarlato V., Masignani V., Giuliani M.M., Arico' B.,
RA   Comanducci M., Jennings G.T., Baldi L., Bartolini E., Capecchi B.,
RA   Galeotti C.L., Luzzi E., Manetti R., Marchetti E., Mora M., Nuti S.,
RA   Ratti G., Santini L., Savino S., Scarselli M., Storni E., Zuo P.,
RA   Broeker M., Hundt E., Knapp B., Blair E., Mason T., Tettelin H.,
RA   Hood D.W., Jeffries A.C., Saunders N.J., Granoff D.M., Venter C.,
RA   Moxon E.R., Grandi G., Rappuoli R.;
RT   "Identification of Vaccine Candidates Against Serogroup B
RT   Meningococcus by Whole-Genome Sequencing.";
RL   Science 287:1816-1820(2000).
DR   EMBL; AF226378; AAF42527.1; -.
DR   InterPro; IPR005594; YadA.
DR   Pfam; PF03895; YadA; 1.
SQ   SEQUENCE   590 AA;  61661 MW;  8AA476AC300D80C8 CRC64;

  Query Match            96.2%;  Score 2260;  DB 2;  Length 590;
  Best Local Similarity  96.7%;  Pred. No. 6e-91;
  Matches  446;  Conservative    3;  Mismatches   12;  Indels    0;  Gaps    0;

Qy          2 DETGLINVETEKLSFGANGKKVNIISDTKGLNFAKETAGTNGDTTVHLNGIGSTLTDMLL 61
              | | | :| |||||| ||| |||| ||||||||||||||||||||||||||||||||| ||
Db        130 DLTDLTSVGTEKLSFSANGNKVNITSDTKGLNFAKETAGTNGDTTVHLNGIGSTLTDTLL 189

Qy         62 NTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLS 121
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 NTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLS 249

Qy        122 ADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKGKGENGSSTDEGEGLVTAK 181
              |||||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||
Db        250 ADTKTTTVNVESKDNGKRTEVKIGAKTSVIKEKDGKLVTGKGKGENGSSTDEGEGLVTAK 309

Qy        182 EVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTKVTFASGNGTTATVSKDDQGNITVK 241
              |||||||||||||||||||||||||||||||| |||||| ||||||||||||||||||
Db        310 EVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVK 369

Qy        242 YDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEIT 301
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        370 YDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEIT 429

Qy        302 RNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGALNVGSKDANKPVRITNVAPGVKE 361
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        430 RNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGALNVGSKDANKPVRITNVAPGVKE 489

Qy        362 GDVTNVAQLKGVAQNLNNRIDNVNGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYLG 421
              |||||||||||||||||||:|||||||||||||||||||||||||||||||||||| |
Db        490 GDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRG 549

Qy        422 EAGYAIGYSSISAGGNWIIKGTASGNSRGHFGASASVGYQW 462
              |||||||||||| |||||||||||||||||||||||||||
Db        550 EAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 590


RESULT 9
Q9JPR8
ID   Q9JPR8      PRELIMINARY;      PRT;   599 AA.
AC   Q9JPR8;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Outer membrane protein GNA992 (NhhA outer membrane protein).
GN   GNA992 OR NHHA.
OS   Neisseria meningitidis.
OC   Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC   Neisseriaceae; Neisseria.
OX   NCBI_TaxID=487;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NGH38;
RX   MEDLINE=20175756; PubMed=10710308;
RA   Pizza M., Scarlato V., Masignani V., Giuliani M.M., Arico' B.,
RA   Comanducci M., Jennings G.T., Baldi L., Bartolini E., Capecchi B.,
RA   Galeotti C.L., Luzzi E., Manetti R., Marchetti E., Mora M., Nuti S.,
RA   Ratti G., Santini L., Savino S., Scarselli M., Storni E., Zuo P.,
RA   Broeker M., Hundt E., Knapp B., Blair E., Mason T., Tettelin H.,
RA   Hood D.W., Jeffries A.C., Saunders N.J., Granoff D.M., Venter C.,
RA   Moxon E.R., Grandi G., Rappuoli R.;
RT   "Identification of Vaccine Candidates Against Serogroup B
RT   Meningococcus by Whole-Genome Sequencing.";
RL   Science 287:1816-1820(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=H38;
RA   Peak I.R., Srikhanta Y., Dieckelman M., Moxon R., Jennings M.P.;
RT   "Identification and characterization of a gene encoding a novel outer
RT   membrane protein of Neisseria meningitidis.";
RL   Submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF226383; AAF42532.1; -.
DR   EMBL; AF157608; AAK68869.1; -.
DR   InterPro; IPR005594; YadA.
DR   Pfam; PF03895; YadA; 1.
SQ   SEQUENCE   599 AA;  62844 MW;  BBA16EBF53C1970C CRC64;

  Query Match            96.2%;  Score 2260;  DB 2;  Length 599;
  Best Local Similarity  96.5%;  Pred. No. 6.1e-91;
  Matches  445;  Conservative    5;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          2 DETGLINVETEKLSFGANGKKVNIISDTKGLNFAKETAGTNGDTTVHLNGIGSTLTDMLL 61
              | | | :|||||||||||| |||| |||||||||||||||||||||||||||||||| ||
Db        139 DLTDLTSVETEKLSFGANGNKVNITSDTKGLNFAKETAGTNGDTTVHLNGIGSTLTDTLL 198

Qy         62 NTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLS 121
              |||||||||||||||:|||||||||||||||||||||||||||||| |||||||||
Db        199 NTGATTNVTNDNVTDDKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVHTYDTVEFLS 258

Qy        122 ADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKGKGENGSSTDEGEGLVTAK 181
              |||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||
Db        259 ADTKTTTVNVESKDNGKRTEVKIGAKTSVIKEKDGKLVTGKGKGENGSSTDEGEGLVTAK 318

Qy        182 EVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTKVTFASGNGTTATVSKDDQGNITVK 241
              |||||||||||||||||||||||||||||||| |||||| ||||||||||||||||||
Db        319 EVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVK 378

Qy        242 YDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEIT 301
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        379 YDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEIT 438

Qy        302 RNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGALNVGSKDANKPVRITNVAPGVKE 361
              |||||||||||||||||||||||||||||||||||||:|||||||||||||||||||||
Db        439 RNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDKGALNVGSKDANKPVRITNVAPGVKE 498

Qy        362 GDVTNVAQLKGVAQNLNNRIDNVNGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYLG 421
              |||||||||||||||||||:|||||||||||||||||||||||||||||||||||| |
Db        499 GDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRG 558

Qy        422 EAGYAIGYSSISAGGNWIIKGTASGNSRGHFGASASVGYQW 462
              |||||||||||| |||||||||||||||||||||||||||
Db        559 EAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 599
```

```
RESULT 10
Q9JPH7
ID   Q9JPH7      PRELIMINARY;      PRT;   594 AA.
AC   Q9JPH7;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Outer membrane protein GNA992 (NhhA outer membrane protein).
GN   GNA992 OR NHHA.
OS   Neisseria meningitidis.
OC   Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC   Neisseriaceae; Neisseria.
OX   NCBI_TaxID=487;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BZ198, and 297-0;
RX   MEDLINE=20175756; PubMed=10710308;
RA   Pizza M., Scarlato V., Masignani V., Giuliani M.M., Arico' B.,
RA   Comanducci M., Jennings G.T., Baldi L., Bartolini E., Capecchi B.,
RA   Galeotti C.L., Luzzi E., Manetti R., Marchetti E., Mora M., Nuti S.,
RA   Ratti G., Santini L., Savino S., Scarselli M., Storni E., Zuo P.,
RA   Broeker M., Hundt E., Knapp B., Blair E., Mason T., Tettelin H.,
RA   Hood D.W., Jeffries A.C., Saunders N.J., Granoff D.M., Venter C.,
RA   Moxon E.R., Grandi G., Rappuoli R.;
RT   "Identification of Vaccine Candidates Against Serogroup B
RT   Meningococcus by Whole-Genome Sequencing.";
RL   Science 287:1816-1820(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BZ198;
RA   Peak I.R., Srikhanta Y., Dieckelman M., Moxon R., Jennings M.P.;
RT   "Identification and characterization of a gene encoding a novel outer
RT   membrane protein of Neisseria meningitidis.";
RL   Submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF226368; AAF42517.1; -.
DR   EMBL; AF226358; AAF42507.1; -.
DR   EMBL; AF157604; AAK68865.1; -.
DR   InterPro; IPR005594; YadA.
DR   Pfam; PF03895; YadA; 1.
SQ   SEQUENCE   594 AA;  62361 MW;  436BDDED68263C5C CRC64;

  Query Match             95.6%;  Score 2247;  DB 2;  Length 594;
  Best Local Similarity   96.3%;  Pred. No. 2.2e-90;
  Matches  444;  Conservative    2;  Mismatches   15;  Indels    0;  Gaps    0;

Qy          2 DETGLINVETEKLSFGANGKKVNIISDTKGLNFAKETAGTNGDTTVHLNGIGSTLTDMLL 61
              | | | :||||||||||||| |||| ||||||||||||||||||| |||||||||||| ||
Db        134 DLTDLTSVETEKLSFGANGNKVNITSDTKGLNFAKETAGTNGDPTVHLNGIGSTLTDTLL 193

Qy         62 NTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLS 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        194 NTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLS 253

Qy        122 ADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKGKGENGSSTDEGEGLVTAK 181
              |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        254 ADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKGKDENGSSTDEGEGLVTAK 313

Qy        182 EVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTKVTFASGNGTTATVSKDDQGNITVK 241
              |||||||||||||||||||||||||||||||||| |||||| ||||||||||||||||||
Db        314 EVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVK 373

Qy        242 YDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEIT 301
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        374 YDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEIT 433

Qy        302 RNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGALNVGSKDANKPVRITNVAPGVKE 361
              |||||||||||| |||||||||||||||||||||||||||| ||||||||||||||||||
Db        434 RNGKNIDIATSMAPQFSSVSLGAGADAPTLSVDDEGALNVGSKDTNKPVRITNVAPGVKE 493

Qy        362 GDVTNVAQLKGVAQNLNNRIDNVNGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYLG 421
              |||||||||||||||||||||:|||||||||||||||||||||||||||||||| || |
Db        494 GDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGDTYRG 553

Qy        422 EAGYAIGYSSISAGGNWIIKGTASGNSRGHFGASASVGYQW 462
              |||||||||||| |||||||||||||||||||||||||||
Db        554 EAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 594


RESULT 11
Q9JPR9
ID   Q9JPR9      PRELIMINARY;      PRT;   598 AA.
AC   Q9JPR9;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Outer membrane protein GNA992.
GN   GNA992.
OS   Neisseria meningitidis.
OC   Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC   Neisseriaceae; Neisseria.
OX   NCBI_TaxID=487;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NGH36;
RX   MEDLINE=20175756; PubMed=10710308;
RA   Pizza M., Scarlato V., Masignani V., Giuliani M.M., Arico' B.,
RA   Comanducci M., Jennings G.T., Baldi L., Bartolini E., Capecchi B.,
RA   Galeotti C.L., Luzzi E., Manetti R., Marchetti E., Mora M., Nuti S.,
RA   Ratti G., Santini L., Savino S., Scarselli M., Storni E., Zuo P.,
RA   Broeker M., Hundt E., Knapp B., Blair E., Mason T., Tettelin H.,
RA   Hood D.W., Jeffries A.C., Saunders N.J., Granoff D.M., Venter C.,
RA   Moxon E.R., Grandi G., Rappuoli R.;
RT   "Identification of Vaccine Candidates Against Serogroup B
RT   Meningococcus by Whole-Genome Sequencing.";
RL   Science 287:1816-1820(2000).
DR   EMBL; AF226382; AAF42531.1; -.
DR   InterPro; IPR005594; YadA.
DR   Pfam; PF03895; YadA; 1.
SQ   SEQUENCE   598 AA;  62718 MW;  9095F8E31AD7C76D CRC64;

  Query Match             95.6%;  Score 2247;  DB 2;  Length 598;
  Best Local Similarity   96.3%;  Pred. No. 2.2e-90;
  Matches  444;  Conservative    2;  Mismatches   15;  Indels    0;  Gaps    0;

Qy          2 DETGLINVETEKLSFGANGKKVNIISDTKGLNFAKETAGTNGDTTVHLNGIGSTLTDMLL 61
              | | | :||||||||||||| |||| ||||||||||||||||||| |||||||||||| ||
Db        138 DLTDLTSVETEKLSFGANGNKVNITSDTKGLNFAKETAGTNGDPTVHLNGIGSTLTDTLL 197

Qy         62 NTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLS 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        198 NTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLS 257

Qy        122 ADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKGKGENGSSTDEGEGLVTAK 181
              |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        258 ADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKGKDENGSSTDEGEGLVTAK 317

Qy        182 EVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTKVTFASGNGTTATVSKDDQGNITVK 241
              |||||||||||||||||||||||||||||||||| |||||| ||||||||||||||||||
Db        318 EVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVK 377

Qy        242 YDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEIT 301
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        378 YDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEIT 437

Qy        302 RNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGALNVGSKDANKPVRITNVAPGVKE 361
              |||||||||||| |||||||||||||||||||||||||||| ||||||||||||||||||
Db        438 RNGKNIDIATSMAPQFSSVSLGAGADAPTLSVDDEGALNVGSKDTNKPVRITNVAPGVKE 497

Qy        362 GDVTNVAQLKGVAQNLNNRIDNVNGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYLG 421
              |||||||||||||||||||||:|||||||||||||||||||||||||||||||| || |
Db        498 GDVTNVAQLKGVAQNLNNRIDNVDGNALAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRG 557

Qy        422 EAGYAIGYSSISAGGNWIIKGTASGNSRGHFGASASVGYQW 462
              |||||||||||| |||||||||||||||||||||||||||
Db        558 EAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 598
```

```
RESULT 12
Q9JPS6
ID   Q9JPS6      PRELIMINARY;      PRT;   600 AA.
AC   Q9JPS6;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Outer membrane protein GNA992.
GN   GNA992.
OS   Neisseria meningitidis.
OC   Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC   Neisseriaceae; Neisseria.
OX   NCBI_TaxID=487;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=E26;
RX   MEDLINE=20175756; PubMed=10710308;
RA   Pizza M., Scarlato V., Masignani V., Giuliani M.M., Arico' B.,
RA   Comanducci M., Jennings G.T., Baldi L., Bartolini E., Capecchi B.,
RA   Galeotti C.L., Luzzi E., Manetti R., Marchetti E., Mora M., Nuti S.,
RA   Ratti G., Santini L., Savino S., Scarselli M., Storni E., Zuo P.,
RA   Broeker M., Hundt E., Knapp B., Blair E., Mason T., Tettelin H.,
RA   Hood D.W., Jeffries A.C., Saunders N.J., Granoff D.M., Venter C.,
RA   Moxon E.R., Grandi G., Rappuoli R.;
RT   "Identification of Vaccine Candidates Against Serogroup B
RT   Meningococcus by Whole-Genome Sequencing.";
RL   Science 287:1816-1820(2000).
DR   EMBL; AF226371; AAF42520.1; -.
DR   InterPro; IPR005594; YadA.
DR   Pfam; PF03895; YadA; 1.
SQ   SEQUENCE   600 AA;  62762 MW;  36256963E0598CD1 CRC64;

  Query Match             94.8%;  Score 2228.5;  DB 2;  Length 600;
  Best Local Similarity   96.3%;  Pred. No. 1.4e-89;
  Matches  442;  Conservative    3;  Mismatches   13;  Indels    1;  Gaps    1;

Qy          4 TGLINVETEKLSFGANGKKVNIISDTKGLNFAKETAGTNGDTTVHLNGIGSTLTDMLLNT 63
              | | :| |||||||||| |||| ||||||||||||||||||||||||||||||||| ||||
Db        143 TDLTSVGTEKLSFGANGNKVNITSDTKGLNFAKETAGTNGDTTVHLNGIGSTLTDTLLNT 202

Qy         64 GATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSAD 123
              |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db        203 GATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVHTYDTVEFLSAD 262

Qy        124 TKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKGKGENGSSTDEGEGLVTAKEV 183
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        263 TKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKGKGENGSSTDEGEGLVTAKEV 322

Qy        184 IDAVNKAGWRMKTTTANGQTGQADKFETVTSGTKVTFASGNGTTATVSKDDQGNITVKYD 243
              |||||||||||||||| |||||||||||| |||||| ||||||||||||||||||||||
Db        323 IDAVNKAGWRMKTTTANGQKGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVKYD 382

Qy        244 VNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRN 303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        383 VNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRN 442

Qy        304 GKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGALNVGSKDANKPVRITNVAPGVKEGD 363
              |||||||||||||||||||||||||||||| : |||||||||||||||||||||||||||
Db        443 GKNIDIATSMTPQFSSVSLGAGADAPTLSVDGD-ALNVGSKDANKPVRITNVAPGVKEGD 501

Qy        364 VTNVAQLKGVAQNLNNRIDNVNGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYLGEA 423
              ||||||||||||||||||:|||||||||||||||||||||||||||||||||||| |||
Db        502 VTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEA 561

Qy        424 GYAIGYSSISAGGNWIIKGTASGNSRGHFGASASVGYQW 462
              |||||||||| |||||||||||||||||||||||||||
Db        562 GYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 600


RESULT 13
Q9JPS4
ID   Q9JPS4      PRELIMINARY;      PRT;   526 AA.
AC   Q9JPS4;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Outer membrane protein GNA992.
GN   GNA992.
OS   Neisseria meningitidis.
OC   Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC   Neisseriaceae; Neisseria.
OX   NCBI_TaxID=487;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NG6/88;
RX   MEDLINE=20175756; PubMed=10710308;
RA   Pizza M., Scarlato V., Masignani V., Giuliani M.M., Arico' B.,
RA   Comanducci M., Jennings G.T., Baldi L., Bartolini E., Capecchi B.,
RA   Galeotti C.L., Luzzi E., Manetti R., Marchetti E., Mora M., Nuti S.,
RA   Ratti G., Santini L., Savino S., Scarselli M., Storni E., Zuo P.,
RA   Broeker M., Hundt E., Knapp B., Blair E., Mason T., Tettelin H.,
RA   Hood D.W., Jeffries A.C., Saunders N.J., Granoff D.M., Venter C.,
RA   Moxon E.R., Grandi G., Rappuoli R.;
RT   "Identification of Vaccine Candidates Against Serogroup B
RT   Meningococcus by Whole-Genome Sequencing.";
RL   Science 287:1816-1820(2000).
DR   EMBL; AF226377; AAF42526.1; -.
DR   InterPro; IPR005594; YadA.
DR   Pfam; PF03895; YadA; 1.
SQ   SEQUENCE   526 AA;  54732 MW;  F543C4234AF807CA CRC64;

  Query Match             94.7%;  Score 2224.5;  DB 2;  Length 526;
  Best Local Similarity   95.7%;  Pred. No. 1.8e-89;
  Matches  441;  Conservative    5;  Mismatches   14;  Indels    1;  Gaps    1;

Qy          2 DETGLINVETEKLSFGANGKKVNIISDTKGLNFAKETAGTNGDTTVHLNGIGSTLTDMLL 61
              | | | :|||||||||||| |||| ||||||||||||||||||||||||||||||| ||
Db         67 DLTDLTSVETEKLSFGANGNKVNITSDTKGLNFAKETAGTNGDTTVHLNGIGSTLTDTLL 126

Qy         62 NTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLS 121
              |||||||||||||||:|||||||||||||||||||||||||||||||| |||||||||
Db        127 NTGATTNVTNDNVTDDKKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVHTYDTVEFLS 186

Qy        122 ADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKGKGENGSSTDEGEGLVTAK 181
              |||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||
Db        187 ADTKTTTVNVESKDNGKRTEVKIGAKTSVIKEKDGKLVTGKGKGENGSSTDEGEGLVTAK 246

Qy        182 EVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTKVTFASGNGTTATVSKDDQGNITVK 241
              |||||||||||||||||||||||||||||||||||||| |||||| ||||||||||||||
Db        247 EVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVK 306

Qy        242 YDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEIT 301
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        307 YDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEIT 366

Qy        302 RNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGALNVGSKDANKPVRITNVAPGVKE 361
              |||||||||||||||||||||||||||||||||||||| : ||||||| ||||||||||||
Db        367 RNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGD-ALNVGSKKDNKPVRITNVAPGVKE 425

Qy        362 GDVTNVAQLKGVAQNLNNRIDNVNGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYLG 421
              |||||||||||||||||||:|||||||||||||||||||||||||||||||||||||| |
Db        426 GDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRG 485

Qy        422 EAGYAIGYSSISAGGNWIIKGTASGNSRGHFGASASVGYQW 462
              |||||||||||| |||||||||||||||||||||||||||
Db        486 EAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 526


RESULT 14
Q9JPS1
ID   Q9JPS1      PRELIMINARY;      PRT;   530 AA.
AC   Q9JPS1;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
```

```
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Outer membrane protein GNA992.
GN   GNA992.
OS   Neisseria meningitidis.
OC   Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC   Neisseriaceae; Neisseria.
OX   NCBI_TaxID=487;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NGF26;
RX   MEDLINE=20175756; PubMed=10710308;
RA   Pizza M., Scarlato V., Masignani V., Giuliani M.M., Arico' B.,
RA   Comanducci M., Jennings G.T., Baldi L., Bartolini E., Capecchi B.,
RA   Galeotti C.L., Luzzi E., Manetti R., Marchetti E., Mora M., Nuti S.,
RA   Ratti G., Santini L., Savino S., Scarselli M., Storni E., Zuo P.,
RA   Broeker M., Hundt E., Knapp B., Blair E., Mason T., Tettelin H.,
RA   Hood D.W., Jeffries A.C., Saunders N.J., Granoff D.M., Venter C.,
RA   Moxon E.R., Grandi G., Rappuoli R.;
RT   "Identification of Vaccine Candidates Against Serogroup B
RT   Meningococcus by Whole-Genome Sequencing.";
RL   Science 287:1816-1820(2000).
DR   EMBL; AF226380; AAF42529.1; -.
DR   InterPro; IPR005594; YadA.
DR   Pfam; PF03895; YadA; 1.
SQ   SEQUENCE   530 AA;  55190 MW;  1F836CA57598515B CRC64;

  Query Match             94.7%;  Score 2224.5;  DB 2;  Length 530;
  Best Local Similarity   95.7%;  Pred. No. 1.8e-89;
  Matches  441;  Conservative    5;  Mismatches   14;  Indels    1;  Gaps    1;

Qy          2 DETGLINVETEKLSFGANGKKVNIISDTKGLNFAKETAGTNGDTTVHLNGIGSTLTDMLL 61
              | | | :|||||||||||| |||| ||||||||||||||||||||||||||||||||| ||
Db         71 DLTDLTSVETEKLSFGANGNKVNITSDTKGLNFAKETAGTNGDTTVHLNGIGSTLTDTLL 130

Qy         62 NTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLS 121
              |||||||||||||||:|||||||||||||||||||||||||||||||| ||||||||
Db        131 NTGATTNVTNDNVTDDKKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVHTYDTVEFLS 190

Qy        122 ADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKGKGENGSSTDEGEGLVTAK 181
              ||||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||
Db        191 ADTKTTTVNVESKDNGKRTEVKIGAKTSVIKEKDGKLVTGKGKGENGSSTDEGEGLVTAK 250

Qy        182 EVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTKVTFASGNGTTATVSKDDQGNITVK 241
              |||||||||||||||||||||||||||||||| |||||| ||||||||||||||||
Db        251 EVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVK 310

Qy        242 YDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEIT 301
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        311 YDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEIT 370

Qy        302 RNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGALNVGSKDANKPVRITNVAPGVKE 361
              |||||||||||||||||||||||||||||||| : ||||||| ||||||||||||||
Db        371 RNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGD-ALNVGSKKDNKPVRITNVAPGVKE 429

Qy        362 GDVTNVAQLKGVAQNLNNRIDNVNGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYLG 421
              |||||||||||||||||||||:|||||||||||||||||||||||||||||||||||| |
Db        430 GDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRG 489

Qy        422 EAGYAIGYSSISAGGNWIIKGTASGNSRGHFGASASVGYQW 462
              |||||||||||| |||||||||||||||||||||||||||
Db        490 EAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 530


RESULT 15
Q93QY4
ID   Q93QY4      PRELIMINARY;      PRT;   594 AA.
AC   Q93QY4;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   NhhA outer membrane protein.
GN   NHHA.
OS   Neisseria meningitidis.
OC   Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC   Neisseriaceae; Neisseria.
OX   NCBI_TaxID=487;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=EG327;
RA   Peak I.R., Srikhanta Y., Dieckelman M., Moxon R., Jennings M.P.;
RT   "Identification and characterization of a gene encoding a novel outer
RT   membrane protein of Neisseria meningitidis.";
RL   Submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF157605; AAK68866.1; -.
DR   InterPro; IPR005594; YadA.
DR   Pfam; PF03895; YadA; 1.
SQ   SEQUENCE   594 AA;  62297 MW;  9DDD48B04B3A8EA2 CRC64;

  Query Match             94.2%;  Score 2213;  DB 2;  Length 594;
  Best Local Similarity   95.0%;  Pred. No. 6.7e-89;
  Matches  438;  Conservative    5;  Mismatches   18;  Indels    0;  Gaps    0;

Qy          2 DETGLINVETEKLSFGANGKKVNIISDTKGLNFAKETAGTNGDTTVHLNGIGSTLTDMLL 61
              | | | :| |||||| ||  |||| ||||||||||:|| ||||||||||||||||| ||
Db        134 DLTDLTSVGTEKLSFSANSNKVNITSDTKGLNFAKKTAETNGDTTVHLNGIGSTLTDTLL 193

Qy         62 NTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLS 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        194 NTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLS 253

Qy        122 ADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKGKGENGSSTDEGEGLVTAK 181
              ||||||||||||||||:|||||||||||||||||||| |||| ||||:||||||||
Db        254 ADTKTTTVNVESKDNGKRTEVKIGAKTSVIKEKDGKLVTGKDKGENDSSTDKGEGLVTAK 313

Qy        182 EVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTKVTFASGNGTTATVSKDDQGNITVK 241
              |||||||||||||||||||||||||||||||| |||||| ||||||||||||||||
Db        314 EVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVM 373

Qy        242 YDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEIT 301
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        374 YDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEIT 433

Qy        302 RNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGALNVGSKDANKPVRITNVAPGVKE 361
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        434 RNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGALNVGSKDANKPVRITNVAPGVKE 493

Qy        362 GDVTNVAQLKGVAQNLNNRIDNVNGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYLG 421
              |||||||||||||||| ||||:|||||||||||||||||||||||||||||||||| |
Db        494 GDVTNVAQLKGVAQNLNNHIDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRG 553

Qy        422 EAGYAIGYSSISAGGNWIIKGTASGNSRGHFGASASVGYQW 462
              |||||||||||| |||||||||||||||||||||||||||
Db        554 EAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 594


Search completed: October  6, 2003, 09:30:51
Job time : 34.9322 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: October 6, 2003, 09:05:45 ; Search time 30.5718 Seconds
(without alignments)
1848.329 Million cell updates/sec

Title: US-09-771-382-37
Perfect score: 1813
Sequence: 1 NNVDFVRTYDTVEFLSADTK.....TASGNSRGHFGASASVGYQW 356

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03:*
1: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1813	100.0	407	22	AAU06184	N. meningitidis PM
2	1808	99.7	433	22	AAU06185	N. meningitidis PM
3	1808	99.7	512	22	AAU06182	N. meningitidis PM
4	1808	99.7	591	20	AAY27202	Amino acid sequenc
5	1808	99.7	591	20	AAY23741	A surface protein
6	1808	99.7	591	20	AAY23746	A surface protein
7	1808	99.7	591	22	AAU06171	N. meningitidis PM
8	1808	99.7	591	22	AAU06175	N. meningitidis EG
9	1807	99.7	502	22	AAU06186	N. meningitidis PM
10	1805	99.6	591	21	AAY57045	BASB029 amino acid
11	1804	99.5	592	20	AAY23737	A surface protein
12	1753.5	96.7	594	20	AAY23740	A surface protein
13	1753.5	96.7	594	21	AAY57044	BASB029 amino acid
14	1753.5	96.7	594	22	AAU06174	N. meningitidis EG
15	1752.5	96.7	592	22	AAU06180	N. meningitidis Z2
16	1748.5	96.4	599	20	AAY23743	A surface protein
17	1748.5	96.4	599	22	AAU06176	N. meningitidis H3
18	1741.5	96.1	594	20	AAY23739	A surface protein
19	1741.5	96.1	594	22	AAU06179	N. meningitidis BZ
20	1736.5	95.8	592	20	AAY27203	Amino acid sequenc
21	1728.5	95.3	513	22	AAU06183	N. meningitidis H4
22	1728.5	95.3	592	20	AAY23744	A surface protein
23	1728.5	95.3	592	22	AAU06172	N. meningitidis H4
24	1728.5	95.3	598	20	AAY23738	A surface protein
25	1728.5	95.3	598	20	AAY23742	A surface protein
26	1728.5	95.3	598	22	AAU06177	N. meningitidis H1
27	1728.5	95.3	598	22	AAU06178	N. meningitidis BZ
28	1715.5	94.6	589	20	AAY23745	A surface protein
29	1715.5	94.6	589	22	AAU06173	N. meningitidis P2
30	1647.5	90.9	604	22	AAU06181	N. meningitidis su
31	608	33.5	1098	17	AAR99392	Haemophilus adhesi
32	600	33.1	116	21	AAB37832	Neisserial conserv
33	581.5	32.1	1094	21	AAB23858	Haemophilus influe
34	569	31.4	2353	17	AAR99393	Haemophilus adhesi
35	569	31.4	2411	21	AAB23860	Haemophilus influe
36	412.5	22.8	679	17	AAR99394	Haemophilus adhesi
37	412.5	22.8	679	21	AAB23855	Haemophilus influe
38	366	20.2	72	21	AAB37830	Neisserial conserv
39	339	18.7	1004	21	AAB23857	Haemophilus influe
40	331.5	18.3	1002	21	AAB23854	Haemophilus influe
41	325	17.9	1104	21	AAB23856	Haemophilus influe
42	325	17.9	1104	21	AAB23859	Haemophilus influe
43	318	17.5	1778	22	ABB52677	Escherichia coli p
44	314	17.3	2139	24	ABP71294	M. catarrhalis sur
45	312.5	17.2	2314	22	AAB69136	M. catarrhalis les

ALIGNMENTS

RESULT 1
AAU06184
ID AAU06184 standard; Protein; 407 AA.
XX
AC AAU06184;
XX
DT 24-OCT-2001 (first entry)
XX
DE N. meningitidis PMC21 NhhA deletion mutant #2.
XX
KW Surface antigen NhhA; meningococcal disease; meningitis vaccine;
KW mutant; mutein.
XX
OS Neisseria meningitidis strain PMC21.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Peptide 1..51
FT /label= Signal_peptide
FT Protein 52..407
FT /label= Mature_NhhA_deletion_mutant_#2
FT /note= "Predicted mature protein, specifically
FT claimed in claim 12"
XX
PN WO200155182-A1.
XX
PD 02-AUG-2001.
XX
PF 25-JAN-2001; 2001WO-AU00069.
XX
PR 25-JAN-2000; 2000US-0177917.
XX

```
PA   (UYQU ) UNIV QUEENSLAND.
XX
PI   Peak IRA,  Jennings MP;
XX
DR   WPI; 2001-488774/53.
DR   N-PSDB; AAS09174.
XX
PT   New NhhA surface antigen polypeptides and polynucleotides from
PT   Neisseria meningitidis, useful in producing vaccines for treating or
PT   preventing broad spectrum of Neisseria meningitidis -
XX
PS   Claim 12; Fig 7; 91pp; English.
XX
CC   The present invention relates to the isolation of novel Neisseria
CC   meningitidis mutant polypeptides of the surface antigen NhhA
CC   (AAU06182-AAU06186). The modified or mutant NhhA polypeptides are
CC   characterised by deletions of non-conserved amino acids, particularly
CC   the deletion of variable regions. The deletion mutants are useful in
CC   diagnostics, therapeutic and prophylactic vaccines against a broader
CC   spectrum of N. meningitidis, and in designing and/or screening of
CC   medicaments. The mutant proteins when used as a vaccine can effectively
CC   immunise against a broader spectrum of N. meningitidis strains than
CC   would be expected from a corresponding wild-type surface antigen.
CC   The present sequence represents N. meningitidis strain PMC21 surface
CC   antigen NhhA deletion mutant #2.
XX
SQ   Sequence   407 AA;

  Query Match             100.0%;  Score 1813;  DB 22;  Length 407;
  Best Local Similarity   100.0%;  Pred. No. 9.1e-118;
  Matches  356;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 NNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKDKG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         52 NNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKDKG 111

Qy         61 ENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        112 ENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKG 171

Qy        121 TTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGK 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        172 TTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGK 231

Qy        181 MDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGDALNVGSKKD 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        232 MDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGDALNVGSKKD 291

Qy        241 NKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAYL 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        292 NKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAYL 351

Qy        301 PGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 356
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        352 PGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 407


RESULT 2
AAU06185
ID   AAU06185 standard; Protein; 433 AA.
XX
AC   AAU06185;
XX
DT   24-OCT-2001  (first entry)
XX
DE   N. meningitidis PMC21 NhhA deletion mutant #3.
XX
KW   Surface antigen NhhA; meningococcal disease; meningitis vaccine;
KW   mutant; mutein.
XX
OS   Neisseria meningitidis strain PMC21.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..51
FT                   /label= Signal_peptide
FT   Protein         52..433
FT                   /label= Mature_NhhA_deletion_mutant_#3
FT                   /note= "Predicted mature protein, specifically
FT                         claimed in claim 12"
XX
PN   WO200155182-A1.
XX
PD   02-AUG-2001.
XX
PF   25-JAN-2001; 2001WO-AU00069.
XX
PR   25-JAN-2000; 2000US-0177917.
XX
PA   (UYQU ) UNIV QUEENSLAND.
XX
PI   Peak IRA,  Jennings MP;
XX
DR   WPI; 2001-488774/53.
DR   N-PSDB; AAS09175.
XX
PT   New NhhA surface antigen polypeptides and polynucleotides from
PT   Neisseria meningitidis, useful in producing vaccines for treating or
PT   preventing broad spectrum of Neisseria meningitidis -
XX
PS   Claim 12; Fig 8; 91pp; English.
XX
CC   The present invention relates to the isolation of novel Neisseria
CC   meningitidis mutant polypeptides of the surface antigen NhhA
CC   (AAU06182-AAU06186). The modified or mutant NhhA polypeptides are
CC   characterised by deletions of non-conserved amino acids, particularly
CC   the deletion of variable regions. The deletion mutants are useful in
CC   diagnostics, therapeutic and prophylactic vaccines against a broader
CC   spectrum of N. meningitidis, and in designing and/or screening of
CC   medicaments. The mutant proteins when used as a vaccine can effectively
CC   immunise against a broader spectrum of N. meningitidis strains than
CC   would be expected from a corresponding wild-type surface antigen.
CC   The present sequence represents N. meningitidis strain PMC21 surface
CC   antigen NhhA deletion mutant #3.
XX
SQ   Sequence   433 AA;

  Query Match             99.7%;  Score 1808;  DB 22;  Length 433;
  Best Local Similarity   99.7%;  Pred. No. 2.2e-117;
  Matches  355;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 NNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKDKG 60
              :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         78 DNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKDKG 137

Qy         61 ENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        138 ENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKG 197

Qy        121 TTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGK 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        198 TTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGK 257

Qy        181 MDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGDALNVGSKKD 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        258 MDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGDALNVGSKKD 317

Qy        241 NKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAYL 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        318 NKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAYL 377

Qy        301 PGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 356
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        378 PGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 433
```

```
RESULT 3
AAU06182
ID   AAU06182 standard; Protein; 512 AA.
XX
AC   AAU06182;
XX
DT   24-OCT-2001  (first entry)
XX
DE   N. meningitidis PMC21 NhhA deletion mutant #1.
XX
KW   Surface antigen NhhA; meningococcal disease; meningitis vaccine;
KW   mutant; mutein.
XX
OS   Neisseria meningitidis strain PMC21.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..51
FT                   /label= Signal_peptide
FT   Protein         52..512
FT                   /label= Mature_NhhA_deletion_mutant_#1
FT                   /note= "Predicted mature protein, specifically
FT                         claimed in claim 12"
XX
PN   WO200155182-A1.
XX
PD   02-AUG-2001.
XX
PF   25-JAN-2001; 2001WO-AU00069.
XX
PR   25-JAN-2000; 2000US-0177917.
XX
PA   (UYQU ) UNIV QUEENSLAND.
XX
PI   Peak IRA,  Jennings MP;
XX
DR   WPI; 2001-488774/53.
DR   N-PSDB; AAS09172.
XX
PT   New NhhA surface antigen polypeptides and polynucleotides from
PT   Neisseria meningitidis, useful in producing vaccines for treating or
PT   preventing broad spectrum of Neisseria meningitidis -
XX
PS   Claim 12; Fig 5; 91pp; English.
XX
CC   The present invention relates to the isolation of novel Neisseria
CC   meningitidis mutant polypeptides of the surface antigen NhhA
CC   (AAU06182-AAU06186). The modified or mutant NhhA polypeptides are
CC   characterised by deletions of non-conserved amino acids, particularly
CC   the deletion of variable regions. The deletion mutants are useful in
CC   diagnostics, therapeutic and prophylactic vaccines against a broader
CC   spectrum of N. meningitidis, and in designing and/or screening of
CC   medicaments. The mutant proteins when used as a vaccine can effectively
CC   immunise against a broader spectrum of N. meningitidis strains than
CC   would be expected from a corresponding wild-type surface antigen.
CC   The present sequence represents N. meningitidis strain PMC21 surface
CC   antigen NhhA deletion mutant #1.
XX
SQ   Sequence   512 AA;

  Query Match             99.7%;  Score 1808;  DB 22;  Length 512;
  Best Local Similarity   99.7%;  Pred. No. 2.7e-117;
  Matches  355;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 NNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKDKG 60
              :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        157 DNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKDKG 216

Qy         61 ENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        217 ENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKG 276

Qy        121 TTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGK 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        277 TTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGK 336

Qy        181 MDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGDALNVGSKKD 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        337 MDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGDALNVGSKKD 396

Qy        241 NKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAYL 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        397 NKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAYL 456

Qy        301 PGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 356
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        457 PGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 512


RESULT 4
AAY27202
ID   AAY27202 standard; Protein; 591 AA.
XX
AC   AAY27202;
XX
DT   24-SEP-1999  (first entry)
XX
DE   Amino acid sequence of N. meningitidis protein ORF40-1.
XX
KW   Neisseria meningitidis protein; pharmaceutical; vaccine; diagnosis;
KW   bacterial infection; treatment.
XX
OS   Neisseria meningitidis.
XX
PN   WO9936544-A2.
XX
PD   22-JUL-1999.
XX
PF   14-JAN-1999;   99WO-IB00103.
XX
PR   09-OCT-1998;   98GB-0022143.
PR   14-JAN-1998;   98GB-0000760.
PR   01-SEP-1998;   98GB-0019015.
XX
PA   (CHIR-) CHIRON SPA.
XX
PI   Grandi G,  Masignani V,  Pizza M,  Rappuoli R,  Scarlato V;
XX
DR   WPI; 1999-444400/37.
DR   N-PSDB; AAX99124.
XX
PT   New protein and its nucleotide sequence, useful in vaccines or
PT   diagnostic compositions for treating and/or preventing Neisseria
PT   meningitidis infections
XX
PS   Claim 1; Page 62; 123pp; English.
XX
CC   The invention provides proteins (AAY27201-245) from Neisseria
CC   meningitidis (strains A and B) and nucleic acid sequences (AAX99123-167)
CC   encoding the proteins. Compositions comprising the protein, nucleic acid
CC   or antibody specific to the protein are useful as pharmaceuticals, e.g. a
CC   vaccine composition or a diagnostic composition. The composition is also
CC   useful for treating or preventing an infection due to Neisserial
CC   bacteria, especially Neisseria meningitidis.
XX
SQ   Sequence   591 AA;

  Query Match             99.7%;  Score 1808;  DB 20;  Length 591;
  Best Local Similarity   99.7%;  Pred. No. 3.2e-117;
  Matches  355;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 NNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKDKG 60
              :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        236 DNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKDKG 295

Qy         61 ENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKG 120
```

```
Db        296 ENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKG 355

Qy        121 TTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGK 180

Db        356 TTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGK 415

Qy        181 MDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGDALNVGSKKD 240

Db        416 MDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGDALNVGSKKD 475

Qy        241 NKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAYL 300

Db        476 NKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAYL 535

Qy        301 PGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 356

Db        536 PGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 591
```

RESULT 5
AAY23741
ID AAY23741 standard; Protein; 591 AA.
XX
AC AAY23741;
XX
DT 08-SEP-1999 (first entry)
XX
DE A surface protein of Neisseria meningitidis.
XX
KW Surface protein; surface glycoprotein; infection; vaccine;
KW immunoreactive peptide.
XX
OS Neisseria meningitidis.
XX
PN WO9931132-A1.
XX
PD 24-JUN-1999.
XX
PF 14-DEC-1998; 98WO-AU01031.
XX
PR 12-DEC-1997; 97GB-0026398.
XX
PA (ISIS-) ISIS INNOVATION LTD.
PA (UYQU) UNIV QUEENSLAND.
XX
PI Jennings MP, Moxon ER, Peak IRA;
XX
DR WPI; 1999-418754/35.
DR N-PSDB; AAX85793.
XX
PT Neisseria meningitidis surface proteins useful for treating N.
PT meningitidis infections
XX
PS Claim 1; Page 104-106; 132pp; English.
XX
CC The present sequence represents a surface protein of Neiserria
CC meningitidis which is approximately 62 kDa. The N. meningitidis
CC surface glycoproteins, nucleic acids, the primers and optionally
CC a thermostable polymerase, or antibodies are useful in a kit for
CC the detection or diagnosis of N. meningitidis infection in humans.
CC The N. meningitidis surface glycoproteins can also be used to
CC prevent or treat N. meningitidis infection in humans, especially
CC in the form of vaccines. The proteins and antibodies can also
CC be used to identify immunoreactive peptides.
XX
SQ Sequence 591 AA;

Query Match 99.7%; Score 1808; DB 20; Length 591;
Best Local Similarity 99.7%; Pred. No. 3.2e-117;
Matches 355; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 NNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKDKG 60

Db        236 DNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKDKG 295

Qy         61 ENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKG 120

Db        296 ENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKG 355

Qy        121 TTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGK 180

Db        356 TTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGK 415

Qy        181 MDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGDALNVGSKKD 240

Db        416 MDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGDALNVGSKKD 475

Qy        241 NKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAYL 300

Db        476 NKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAYL 535

Qy        301 PGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 356

Db        536 PGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 591
```

RESULT 6
AAY23746
ID AAY23746 standard; Protein; 591 AA.
XX
AC AAY23746;
XX
DT 08-SEP-1999 (first entry)
XX
DE A surface protein of Neisseria meningitidis.
XX
KW Surface protein; surface glycoprotein; infection; vaccine;
KW immunoreactive peptide.
XX
OS Neisseria meningitidis.
XX
PN WO9931132-A1.
XX
PD 24-JUN-1999.
XX
PF 14-DEC-1998; 98WO-AU01031.
XX
PR 12-DEC-1997; 97GB-0026398.
XX
PA (ISIS-) ISIS INNOVATION LTD.
PA (UYQU) UNIV QUEENSLAND.
XX
PI Jennings MP, Moxon ER, Peak IRA;
XX
DR WPI; 1999-418754/35.
DR N-PSDB; AAX85798.
XX
PT Neisseria meningitidis surface proteins useful for treating N.
PT meningitidis infections
XX
PS Claim 1; Page 127-128; 132pp; English.
XX
CC The present sequence represents a surface protein of Neiserria
CC meningitidis which is approximately 62 kDa. The N. meningitidis
CC surface glycoproteins, nucleic acids, the primers and optionally
CC a thermostable polymerase, or antibodies are useful in a kit for
CC the detection or diagnosis of N. meningitidis infection in humans.
CC The N. meningitidis surface glycoproteins can also be used to
CC prevent or treat N. meningitidis infection in humans, especially
CC in the form of vaccines. The proteins and antibodies can also
CC be used to identify immunoreactive peptides.
XX
SQ Sequence 591 AA;

Query Match 99.7%; Score 1808; DB 20; Length 591;

```
  Best Local Similarity   99.7%;  Pred. No. 3.2e-117;
  Matches  355;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 NNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKDKG 60
              :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        236 DNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKDKG 295

Qy         61 ENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        296 ENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKG 355

Qy        121 TTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGK 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        356 TTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGK 415

Qy        181 MDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGDALNVGSKKD 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        416 MDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGDALNVGSKKD 475

Qy        241 NKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAYL 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        476 NKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAYL 535

Qy        301 PGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 356
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        536 PGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 591


RESULT 7
AAU06171
ID   AAU06171 standard; Protein; 591 AA.
XX
AC   AAU06171;
XX
DT   24-OCT-2001  (first entry)
XX
DE   N. meningitidis PMC21 surface antigen NhhA polypeptide sequence.
XX
KW   Surface antigen NhhA; meningococcal disease; meningitis vaccine.
XX
OS   Neisseria meningitidis strain PMC21.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..51
FT                   /label= Signal_peptide
FT   Region          1..50
FT                   /label= C1
FT                   /note= "Conserved region 1"
FT   Region          51..108
FT                   /label= V1
FT                   /note= "Variable region 1"
FT   Protein         52..591
FT                   /label= Mature_NhhA
FT                   /note= "Predicted mature protein, specifically
FT                         claimed in claim 12"
FT   Region          109..120
FT                   /label= C2
FT                   /note= "Conserved region 2"
FT   Region          121..124
FT                   /label= V2
FT                   /note= "Variable region 2"
FT   Region          125..188
FT                   /label= C3
FT                   /note= "Conserved region 3"
FT   Region          189..210
FT                   /label= V3
FT                   /note= "Variable region 3"
FT   Region          211..229
FT                   /label= C4
FT                   /note= "Conserved region 4"
FT   Region          230..236
FT                   /label= V4
FT                   /note= "Variable region 4"
FT   Region          237..591
FT                   /label= C5
FT                   /note= "Conserved region 5"
XX
PN   WO200155182-A1.
XX
PD   02-AUG-2001.
XX
PF   25-JAN-2001; 2001WO-AU00069.
XX
PR   25-JAN-2000; 2000US-0177917.
XX
PA   (UYQU ) UNIV QUEENSLAND.
XX
PI   Peak IRA,  Jennings MP;
XX
DR   WPI; 2001-488774/53.
DR   N-PSDB; AAS09161.
XX
PT   New NhhA surface antigen polypeptides and polynucleotides from
PT   Neisseria meningitidis, useful in producing vaccines for treating or
PT   preventing broad spectrum of Neisseria meningitidis -
XX
PS   Claim 9; Fig 1; 91pp; English.
XX
CC   The present invention relates to the isolation of novel Neisseria
CC   meningitidis mutant polypeptides of the surface antigen NhhA
CC   (AAU06182-AAU06186). The modified or mutant NhhA polypeptides are
CC   characterised by deletions of non-conserved amino acids, particularly
CC   the deletion of variable regions. The deletion mutants are useful in
CC   diagnostics, therapeutic and prophylactic vaccines against a broader
CC   spectrum of N. meningitidis, and in designing and/or screening of
CC   medicaments. The mutant proteins when used as a vaccine can effectively
CC   immunise against a broader spectrum of N. meningitidis strains than
CC   would be expected from a corresponding wild-type surface antigen.
CC   The present sequence representing the wild type surface antigen NhhA
CC   from N. meningitidis strain PMC21 is 1 of 10 NhhA polypeptide sequences
CC   (AAU06171-AAU06180) from 10 different N. meningitidis strains given in
CC   the present invention.
XX
SQ   Sequence   591 AA;

  Query Match             99.7%;  Score 1808;  DB 22;  Length 591;
  Best Local Similarity   99.7%;  Pred. No. 3.2e-117;
  Matches  355;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 NNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKDKG 60
              :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        236 DNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKDKG 295

Qy         61 ENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        296 ENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKG 355

Qy        121 TTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGK 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        356 TTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGK 415

Qy        181 MDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGDALNVGSKKD 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        416 MDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGDALNVGSKKD 475

Qy        241 NKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAYL 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        476 NKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAYL 535

Qy        301 PGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 356
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        536 PGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 591


RESULT 8
AAU06175
```

```
ID   AAU06175 standard; Protein; 591 AA.
XX
AC   AAU06175;
XX
DT   24-OCT-2001  (first entry)
XX
DE   N. meningitidis EG329 surface antigen NhhA polypeptide sequence.
XX
KW   Surface antigen NhhA; meningococcal disease; meningitis vaccine.
XX
OS   Neisseria meningitidis strain EG329.
XX
FH   Key             Location/Qualifiers
FT   Region          1..50
FT                   /label= C1
FT                   /note= "Conserved region 1"
FT   Region          51..108
FT                   /label= V1
FT                   /note= "Variable region 1"
FT   Region          109..120
FT                   /label= C2
FT                   /note= "Conserved region 2"
FT   Region          121..124
FT                   /label= V2
FT                   /note= "Variable region 2"
FT   Region          125..188
FT                   /label= C3
FT                   /note= "Conserved region 3"
FT   Region          189..210
FT                   /label= V3
FT                   /note= "Variable region 3"
FT   Region          211..229
FT                   /label= C4
FT                   /note= "Conserved region 4"
FT   Region          230..236
FT                   /label= V4
FT                   /note= "Variable region 4"
FT   Region          237..591
FT                   /label= C5
FT                   /note= "Conserved region 5"
XX
PN   WO200155182-A1.
XX
PD   02-AUG-2001.
XX
PF   25-JAN-2001; 2001WO-AU00069.
XX
PR   25-JAN-2000; 2000US-0177917.
XX
PA   (UYQU ) UNIV QUEENSLAND.
XX
PI   Peak IRA,  Jennings MP;
XX
DR   WPI; 2001-488774/53.
DR   N-PSDB; AAS09165.
XX
PT   New NhhA surface antigen polypeptides and polynucleotides from
PT   Neisseria meningitidis, useful in producing vaccines for treating or
PT   preventing broad spectrum of Neisseria meningitidis -
XX
PS   Claim 9; Fig 1; 91pp; English.
XX
CC   The present invention relates to the isolation of novel Neisseria
CC   meningitidis mutant polypeptides of the surface antigen NhhA
CC   (AAU06182-AAU06186). The modified or mutant NhhA polypeptides are
CC   characterised by deletions of non-conserved amino acids, particularly
CC   the deletion of variable regions. The deletion mutants are useful in
CC   diagnostics, therapeutic and prophylactic vaccines against a broader
CC   spectrum of N. meningitidis, and in designing and/or screening of
CC   medicaments. The mutant proteins when used as a vaccine can effectively
CC   immunise against a broader spectrum of N. meningitidis strains than
CC   would be expected from a corresponding wild-type surface antigen.
CC   The present sequence representing the wild type surface antigen NhhA
CC   from N. meningitidis strain EG329 is 1 of 10 NhhA polypeptide sequences
CC   (AAU06171-AAU06180) from 10 different N. meningitidis strains given in
CC   the present invention.
XX
SQ   Sequence   591 AA;

  Query Match             99.7%;  Score 1808;  DB 22;  Length 591;
  Best Local Similarity   99.7%;  Pred. No. 3.2e-117;
  Matches  355;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 NNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKDKG 60
              :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        236 DNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKDKG 295

Qy         61 ENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        296 ENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKG 355

Qy        121 TTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGK 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        356 TTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGK 415

Qy        181 MDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGDALNVGSKKD 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        416 MDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGDALNVGSKKD 475

Qy        241 NKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAYL 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        476 NKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAYL 535

Qy        301 PGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 356
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        536 PGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 591


RESULT 9
AAU06186
ID   AAU06186 standard; Protein; 502 AA.
XX
AC   AAU06186;
XX
DT   24-OCT-2001  (first entry)
XX
DE   N. meningitidis PMC21 NhhA deletion mutant #4.
XX
KW   Surface antigen NhhA; meningococcal disease; meningitis vaccine;
KW   mutant; mutein.
XX
OS   Neisseria meningitidis strain PMC21.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..49
FT                   /label= Signal_peptide
FT   Protein         50..502
FT                   /label= Mature_NhhA_deletion_mutant_#4
FT                   /note= "Predicted mature protein, specifically
FT                         claimed in claim 12"
XX
PN   WO200155182-A1.
XX
PD   02-AUG-2001.
XX
PF   25-JAN-2001; 2001WO-AU00069.
XX
PR   25-JAN-2000; 2000US-0177917.
XX
PA   (UYQU ) UNIV QUEENSLAND.
XX
PI   Peak IRA,  Jennings MP;
XX
DR   WPI; 2001-488774/53.
DR   N-PSDB; AAS09176.
```

```
XX
PT   New NhhA surface antigen polypeptides and polynucleotides from
PT   Neisseria meningitidis, useful in producing vaccines for treating or
PT   preventing broad spectrum of Neisseria meningitidis -
XX
PS   Claim 12; Fig 9; 91pp; English.
XX
CC   The present invention relates to the isolation of novel Neisseria
CC   meningitidis mutant polypeptides of the surface antigen NhhA
CC   (AAU06182-AAU06186). The modified or mutant NhhA polypeptides are
CC   characterised by deletions of non-conserved amino acids, particularly
CC   the deletion of variable regions. The deletion mutants are useful in
CC   diagnostics, therapeutic and prophylactic vaccines against a broader
CC   spectrum of N. meningitidis, and in designing and/or screening of
CC   medicaments. The mutant proteins when used as a vaccine can effectively
CC   immunise against a broader spectrum of N. meningitidis strains than
CC   would be expected from a corresponding wild-type surface antigen.
CC   The present sequence represents N. meningitidis strain PMC21 surface
CC   antigen NhhA deletion mutant #4.
XX
SQ   Sequence   502 AA;

  Query Match             99.7%;  Score 1807;  DB 22;  Length 502;
  Best Local Similarity   100.0%;  Pred. No. 3.1e-117;
  Matches  355;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 NVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKDKGE 61
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        148 NVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKDKGE 207

Qy         62 NGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKGT 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        208 NGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKGT 267

Qy        122 TATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKM 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        268 TATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKM 327

Qy        182 DETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGDALNVGSKKDN 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        328 DETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGDALNVGSKKDN 387

Qy        242 KPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAYLP 301
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        388 KPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAYLP 447

Qy        302 GKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 356
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        448 GKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 502


RESULT 10
AAY57045
ID   AAY57045 standard; Protein; 591 AA.
XX
AC   AAY57045;
XX
DT   21-FEB-2000  (first entry)
XX
DE   BASB029 amino acid sequence from N. meningitidis strain H44/76.
XX
KW   BASB029; Nisseria meningitidis; surface fibril protein; HSF; diagnosis;
KW   infection; treatment; prevent; antibacterial drug.
XX
OS   Neisseria meningitidis.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 90
FT                   /note= "Encoded by AAT"
FT   Misc-difference 92
FT                   /note= "Encoded by GAT"
FT   Misc-difference 98
FT                   /note= "Encoded by AAC"
FT   Misc-difference 108
FT                   /note= "Encoded by AATC"
FT   Misc-difference 123
FT                   /note= "Encoded by ACA"
FT   Misc-difference 269
FT                   /note= "Encoded by AAA"
FT   Misc-difference 389
FT                   /note= "Encoded by CGT"
XX
PN   WO9958683-A2.
XX
PD   18-NOV-1999.
XX
PF   07-MAY-1999;   99WO-EP03255.
XX
PR   13-MAY-1998;   98GB-0010276.
XX
PA   (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PI   Ruelle J;
XX
DR   WPI; 2000-053103/04.
DR   N-PSDB; AAZ39865.
XX
PT   New polypeptide from neisseria meningitidis useful for diagnosis,
PT   treatment or prevention of bacterial infections in mammal  -
XX
PS   Claim 4; Fig 2; 74pp; English.
XX
CC   This is the Nisseria meningitidis BASB029 amino acid sequence from
CC   serogroup B strain H44/76. The BASB029 protein is homologous to the
CC   Haemophilus influenzae surface fibril (HSF) protein. The invention
CC   relates to BASB029 polynucleotide sequences (AAZ39864-Z39865) and
CC   polypeptide sequences (AAY57044-Y57045) and their immunogenic fragments.
CC   BASB029 polypeptides are useful in a method of diagnosing a Neisseria
CC   meningitidis infection in a mammal. Compositions containing BASB029
CC   polynucleotides and polypeptides are useful for generating an immune
CC   response in an animal. A therapeutic composition comprising an antibody
CC   directed against BAsB029 is useful in treating humans with Neisseria
CC   meningitidis disease. The polynucleotide is useful in the diagnosis of
CC   the stage of infection, type of infection, susceptibility to an
CC   infection which results from increased or decreased expression of the
CC   polynucleotide, and for therapeutic or prophylactic purposes,
CC   particularly genetic immunisation. Antibodies against BASB029
CC   polynucleotides and polypeptides are also useful for treating infections
CC   particularly bacterial infections. The protein is useful in the
CC   screening and development of antibacterial drugs. Fused recombinant
CC   protein is useful for the stimulation of the immune system of an organism
CC   receiving the protein.
XX
SQ   Sequence   591 AA;

  Query Match             99.6%;  Score 1805;  DB 21;  Length 591;
  Best Local Similarity   99.4%;  Pred. No. 5.1e-117;
  Matches  354;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 NNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKDKG 60
              :|||||||||||||||||||||||||||||:|||||||||||||||||||||||||
Db        236 DNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKRTEVKIGAKTSVIKEKDGKLVTGKDKG 295

Qy         61 ENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        296 ENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKG 355

Qy        121 TTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGK 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        356 TTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGK 415

Qy        181 MDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGDALNVGSKKD 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        416 MDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGDALNVGSKKD 475

Qy        241 NKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAYL 300
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        476 NKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAYL 535

Qy        301 PGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 356
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        536 PGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 591


RESULT 11
AAY23737
ID   AAY23737 standard; Protein; 592 AA.
XX
AC   AAY23737;
XX
DT   08-SEP-1999  (first entry)
XX
DE   A surface protein of Neisseria meningitidis.
XX
KW   Surface protein; surface glycoprotein; infection; vaccine;
KW   immunoreactive peptide.
XX
OS   Neisseria meningitidis.
XX
PN   WO9931132-A1.
XX
PD   24-JUN-1999.
XX
PF   14-DEC-1998;   98WO-AU01031.
XX
PR   12-DEC-1997;   97GB-0026398.
XX
PA   (ISIS-) ISIS INNOVATION LTD.
PA   (UYQU ) UNIV QUEENSLAND.
XX
PI   Jennings MP,  Moxon ER,  Peak IRA;
XX
DR   WPI; 1999-418754/35.
DR   N-PSDB; AAX85788.
XX
PT   Neisseria meningitidis surface proteins useful for treating N.
PT   meningitidis infections
XX
PS   Claim 1; Page 86-87; 132pp; English.
XX
CC   The present sequence represents a surface protein of Neiserria
CC   meningitidis which is approximately 62 kDa. The N. meningitidis
CC   surface glycoproteins, nucleic acids, the primers and optionally
CC   a thermostable polymerase, or antibodies are useful in a kit for
CC   the detection or diagnosis of N. meningitidis infection in humans.
CC   The N. meningitidis surface glycoproteins can also be used to
CC   prevent or treat N. meningitidis infection in humans, especially
CC   in the form of vaccines. The proteins and antibodies can also
CC   be used to identify immunoreactive peptides.
XX
SQ   Sequence   592 AA;

  Query Match           99.5%;  Score 1804;  DB 20;  Length 592;
  Best Local Similarity   99.4%;  Pred. No. 6.1e-117;
  Matches  354;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 NNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKDKG 60
              :|||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        237 DNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGVKTSVIKEKDGKLVTGKDKG 296

Qy         61 ENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        297 ENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKG 356

Qy        121 TTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGK 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        357 TTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGK 416

Qy        181 MDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGDALNVGSKKD 240
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        417 MDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGDALNVGSKKD 476

Qy        241 NKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAYL 300
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        477 NKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAYL 536

Qy        301 PGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 356
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        537 PGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 592


RESULT 12
AAY23740
ID   AAY23740 standard; Protein; 594 AA.
XX
AC   AAY23740;
XX
DT   08-SEP-1999  (first entry)
XX
DE   A surface protein of Neisseria meningitidis.
XX
KW   Surface protein; surface glycoprotein; infection; vaccine;
KW   immunoreactive peptide.
XX
OS   Neisseria meningitidis.
XX
PN   WO9931132-A1.
XX
PD   24-JUN-1999.
XX
PF   14-DEC-1998;   98WO-AU01031.
XX
PR   12-DEC-1997;   97GB-0026398.
XX
PA   (ISIS-) ISIS INNOVATION LTD.
PA   (UYQU ) UNIV QUEENSLAND.
XX
PI   Jennings MP,  Moxon ER,  Peak IRA;
XX
DR   WPI; 1999-418754/35.
DR   N-PSDB; AAX85792.
XX
PT   Neisseria meningitidis surface proteins useful for treating N.
PT   meningitidis infections
XX
PS   Claim 1; Page 100-101; 132pp; English.
XX
CC   The present sequence represents a surface protein of Neiserria
CC   meningitidis which is approximately 62 kDa. The N. meningitidis
CC   surface glycoproteins, nucleic acids, the primers and optionally
CC   a thermostable polymerase, or antibodies are useful in a kit for
CC   the detection or diagnosis of N. meningitidis infection in humans.
CC   The N. meningitidis surface glycoproteins can also be used to
CC   prevent or treat N. meningitidis infection in humans, especially
CC   in the form of vaccines. The proteins and antibodies can also
CC   be used to identify immunoreactive peptides.
XX
SQ   Sequence   594 AA;

  Query Match           96.7%;  Score 1753.5;  DB 20;  Length 594;
  Best Local Similarity   97.2%;  Pred. No. 1.9e-113;
  Matches  347;  Conservative    4;  Mismatches    5;  Indels    1;  Gaps    1;

Qy          1 NNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKDKG 60
              :|||||||||||||||||||||||||||||:||||||||||||||||||||||||||
Db        238 DNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKRTEVKIGAKTSVIKEKDGKLVTGKDKG 297

Qy         61 ENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKG 120
              || ||||:||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        298 ENDSSTDKGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKG 357

Qy        121 TTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGK 180
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        358 TTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGK 417

Qy        181 MDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGD-ALNVGSKK 239
                 |||||||||||||||||||||||||||||||||||||||||||||||||| : |||||||
Db        418 MDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGALNVGSKD 477

Qy        240 DNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAY 299
                 |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        478 ANKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNHIDNVDGNARAGIAQAIATAGLVQAY 537

Qy        300 LPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 356
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        538 LPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 594


RESULT 13
AAY57044
ID   AAY57044 standard; Protein; 594 AA.
XX
AC   AAY57044;
XX
DT   21-FEB-2000  (first entry)
XX
DE   BASB029 amino acid sequence from N. meningitidis strain ATCC13090.
XX
KW   BASB029; Nisseria meningitidis; surface fibril protein; HSF; diagnosis;
KW   infection; treatment; prevent; antibacterial drug.
XX
OS   Neisseria meningitidis.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 104
FT                   /note= "Encoded by AATC"
XX
PN   WO9958683-A2.
XX
PD   18-NOV-1999.
XX
PF   07-MAY-1999;   99WO-EP03255.
XX
PR   13-MAY-1998;   98GB-0010276.
XX
PA   (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PI   Ruelle J;
XX
DR   WPI; 2000-053103/04.
DR   N-PSDB; AAZ39864.
XX
PT   New polypeptide from neisseria meningitidis useful for diagnosis,
PT   treatment or prevention of bacterial infections in mammal -
XX
PS   Claim 4; Fig 2; 74pp; English.
XX
CC   This is the Nisseria meningitidis BASB029 amino acid sequence from
CC   serogroup B strain ATCC13090. The BASB029 protein is homologous to the
CC   Haemophilus influenzae surface fibril (HSF) protein. The invention
CC   relates to BASB029 polynucleotide sequences (AAZ39864-Z39865) and
CC   polypeptide sequences (AAY57044-Y57045) and their immunogenic fragments.
CC   BASB029 polypeptides are useful in a method of diagnosing a Neisseria
CC   meningitidis infection in a mammal. Compositions containing BASB029
CC   polynucleotides and polypeptides are useful for generating an immune
CC   response in an animal. A therapeutic composition comprising an antibody
CC   directed against BAsB029 is useful in treating humans with Neisseria
CC   meningitidis disease. The polynucleotide is useful in the diagnosis of
CC   the stage of infection, type of infection, susceptibility to an
CC   infection which results from increased or decreased expression of the
CC   polynucleotide, and for therapeutic or prophylactic purposes,
CC   particularly genetic immunisation. Antibodies against BASB029
CC   polynucleotides and polypeptides are also useful for treating infections
CC   particularly bacterial infections. The protein is useful in the
CC   screening and development of antibacterial drugs. Fused recombinant
CC   protein is useful for the stimulation of the immune system of an organism
CC   receiving the protein.
XX
SQ   Sequence   594 AA;

  Query Match             96.7%;  Score 1753.5;  DB 21;  Length 594;
  Best Local Similarity   97.2%;  Pred. No. 1.9e-113;
  Matches  347;  Conservative    4;  Mismatches    5;  Indels    1;  Gaps    1;

Qy          1 NNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKDKG 60
              :|||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||
Db        238 DNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKRTEVKIGAKTSVIKEKDGKLVTGKDKG 297

Qy         61 ENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKG 120
              || ||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        298 ENDSSTDKGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKG 357

Qy        121 TTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGK 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        358 TTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGK 417

Qy        181 MDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGD-ALNVGSKK 239
              |||||||||||||||||||||||||||||||||||||||||||||||||| : |||||||
Db        418 MDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGALNVGSKD 477

Qy        240 DNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAY 299
              |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        478 ANKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNHIDNVDGNARAGIAQAIATAGLVQAY 537

Qy        300 LPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 356
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        538 LPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 594


RESULT 14
AAU06174
ID   AAU06174 standard; Protein; 594 AA.
XX
AC   AAU06174;
XX
DT   24-OCT-2001  (first entry)
XX
DE   N. meningitidis EG327 surface antigen NhhA polypeptide sequence.
XX
KW   Surface antigen NhhA; meningococcal disease; meningitis vaccine.
XX
OS   Neisseria meningitidis strain EG327.
XX
FH   Key             Location/Qualifiers
FT   Region          1..50
FT                   /label= C1
FT                   /note= "Conserved region 1"
FT   Region          51..104
FT                   /label= V1
FT                   /note= "Variable region 1"
FT   Region          105..116
FT                   /label= C2
FT                   /note= "Conserved region 2"
FT   Region          117..126
FT                   /label= V2
FT                   /note= "Variable region 2"
FT   Region          127..190
FT                   /label= C3
FT                   /note= "Conserved region 3"
FT   Region          191..212
FT                   /label= V3
FT                   /note= "Variable region 3"
FT   Region          213..231
FT                   /label= C4
FT                   /note= "Conserved region 4"
FT   Region          232..238
FT                   /label= V4
FT                   /note= "Variable region 4"
```

```
FT   Region          239..594
FT                   /label= C5
FT                   /note= "Conserved region 5"
XX
PN   WO200155182-A1.
XX
PD   02-AUG-2001.
XX
PF   25-JAN-2001; 2001WO-AU00069.
XX
PR   25-JAN-2000; 2000US-0177917.
XX
PA   (UYQU ) UNIV QUEENSLAND.
XX
PI   Peak IRA,  Jennings MP;
XX
DR   WPI; 2001-488774/53.
DR   N-PSDB; AAS09164.
XX
PT   New NhhA surface antigen polypeptides and polynucleotides from
PT   Neisseria meningitidis, useful in producing vaccines for treating or
PT   preventing broad spectrum of Neisseria meningitidis -
XX
PS   Claim 9; Fig 1; 91pp; English.
XX
CC   The present invention relates to the isolation of novel Neisseria
CC   meningitidis mutant polypeptides of the surface antigen NhhA
CC   (AAU06182-AAU06186). The modified or mutant NhhA polypeptides are
CC   characterised by deletions of non-conserved amino acids, particularly
CC   the deletion of variable regions. The deletion mutants are useful in
CC   diagnostics, therapeutic and prophylactic vaccines against a broader
CC   spectrum of N. meningitidis, and in designing and/or screening of
CC   medicaments. The mutant proteins when used as a vaccine can effectively
CC   immunise against a broader spectrum of N. meningitidis strains than
CC   would be expected from a corresponding wild-type surface antigen.
CC   The present sequence representing the wild type surface antigen NhhA
CC   from N. meningitidis strain EG327 is 1 of 10 NhhA polypeptide sequences
CC   (AAU06171-AAU06180) from 10 different N. meningitidis strains given in
CC   the present invention.
XX
SQ   Sequence   594 AA;

  Query Match             96.7%;  Score 1753.5;  DB 22;  Length 594;
  Best Local Similarity   97.2%;  Pred. No. 1.9e-113;
  Matches  347;  Conservative    4;  Mismatches    5;  Indels    1;  Gaps    1;

Qy          1 NNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKDKG 60
              :|||||||||||||||||||||||||||||||:||||||||||||||||||||||||||
Db        238 DNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKRTEVKIGAKTSVIKEKDGKLVTGKDKG 297

Qy         61 ENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKG 120
              || ||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        298 ENDSSTDKGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKG 357

Qy        121 TTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGK 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        358 TTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGK 417

Qy        181 MDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGD-ALNVGSKK 239
              |||||||||||||||||||||||||||||||||||||||||||||||| : |||||||
Db        418 MDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGALNVGSKD 477

Qy        240 DNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAY 299
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        478 ANKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNHIDNVDGNARAGIAQAIATAGLVQAY 537

Qy        300 LPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 356
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        538 LPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 594


RESULT 15
AAU06180
ID   AAU06180 standard; Protein; 592 AA.
XX
AC   AAU06180;
XX
DT   24-OCT-2001  (first entry)
XX
DE   N. meningitidis Z2491 surface antigen NhhA polypeptide sequence.
XX
KW   Surface antigen NhhA; meningococcal disease; meningitis vaccine.
XX
OS   Neisseria meningitidis strain Z2491.
XX
FH   Key             Location/Qualifiers
FT   Region          1..50
FT                   /label= C1
FT                   /note= "Conserved region 1"
FT   Region          51..102
FT                   /label= V1
FT                   /note= "Variable region 1"
FT   Region          103..114
FT                   /label= C2
FT                   /note= "Conserved region 2"
FT   Region          115..124
FT                   /label= V2
FT                   /note= "Variable region 2"
FT   Region          125..188
FT                   /label= C3
FT                   /note= "Conserved region 3"
FT   Region          189..208
FT                   /label= V3
FT                   /note= "Variable region 3"
FT   Region          209..227
FT                   /label= C4
FT                   /note= "Conserved region 4"
FT   Region          228..236
FT                   /label= V4
FT                   /note= "Variable region 4"
FT   Region          237..592
FT                   /label= C5
FT                   /note= "Conserved region 5"
XX
PN   WO200155182-A1.
XX
PD   02-AUG-2001.
XX
PF   25-JAN-2001; 2001WO-AU00069.
XX
PR   25-JAN-2000; 2000US-0177917.
XX
PA   (UYQU ) UNIV QUEENSLAND.
XX
PI   Peak IRA,  Jennings MP;
XX
DR   WPI; 2001-488774/53.
DR   N-PSDB; AAS09170.
XX
PT   New NhhA surface antigen polypeptides and polynucleotides from
PT   Neisseria meningitidis, useful in producing vaccines for treating or
PT   preventing broad spectrum of Neisseria meningitidis -
XX
PS   Claim 9; Fig 1; 91pp; English.
XX
CC   The present invention relates to the isolation of novel Neisseria
CC   meningitidis mutant polypeptides of the surface antigen NhhA
CC   (AAU06182-AAU06186). The modified or mutant NhhA polypeptides are
CC   characterised by deletions of non-conserved amino acids, particularly
CC   the deletion of variable regions. The deletion mutants are useful in
CC   diagnostics, therapeutic and prophylactic vaccines against a broader
CC   spectrum of N. meningitidis, and in designing and/or screening of
CC   medicaments. The mutant proteins when used as a vaccine can effectively
CC   immunise against a broader spectrum of N. meningitidis strains than
CC   would be expected from a corresponding wild-type surface antigen.
CC   The present sequence representing the wild type surface antigen NhhA
```

```
CC   from N. meningitidis strain Z2491 is 1 of 10 NhhA polypeptide sequences
CC   (AAU06171-AAU06180) from 10 different N. meningitidis strains given in
CC   the present invention.
XX
SQ   Sequence   592 AA;

  Query Match             96.7%;  Score 1752.5;  DB 22;  Length 592;
  Best Local Similarity   97.5%;  Pred. No. 2.2e-113;
  Matches  347;  Conservative    3;  Mismatches    5;  Indels    1;  Gaps    1;

Qy          2 NVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKDKGE 61
              ||||||||||||||||||||||||||||||:||||||||||||||||||||| |||
Db        237 NVDFVRTYDTVEFLSADTKTTTVNVESKDNGKRTEVKIGAKTSVIKEKDGKLVTGKGKGE 296

Qy         62 NGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKGT 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        297 NGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKGT 356

Qy        122 TATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKM 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        357 TATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKM 416

Qy        182 DETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGD-ALNVGSKKD 240
              |||||||||||||:||||||||||| |||||||||||||||||| : |||||||
Db        417 DETVNINAGNNIEISRNGKNIDIATSMAPQFSSVSLGAGADAPTLSVDDEGALNVGSKDA 476

Qy        241 NKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAYL 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        477 NKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAYL 536

Qy        301 PGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 356
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        537 PGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 592


Search completed: October  6, 2003, 09:22:43
Job time : 31.5718 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: October 6, 2003, 09:14:45 ; Search time 9.52015 Seconds
(without alignments)
1582.188 Million cell updates/sec

Title: US-09-771-382-37
Perfect score: 1813
Sequence: 1 NNVDFVRTYDTVEFLSADTK.....TASGNSRGHFGASASVGYQW 356

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1808	99.7	591	3	US-09-377-155-11	Sequence 11, Appl
2	1808	99.7	591	3	US-09-377-155-21	Sequence 21, Appl
3	1808	99.7	591	4	US-09-669-974-11	Sequence 11, Appl
4	1808	99.7	591	4	US-09-669-974-21	Sequence 21, Appl
5	1804	99.5	592	3	US-09-377-155-2	Sequence 2, Appli
6	1804	99.5	592	4	US-09-669-974-2	Sequence 2, Appli
7	1753.5	96.7	594	3	US-09-377-155-9	Sequence 9, Appli
8	1753.5	96.7	594	4	US-09-669-974-9	Sequence 9, Appli
9	1748.5	96.4	599	3	US-09-377-155-15	Sequence 15, Appl
10	1748.5	96.4	599	4	US-09-669-974-15	Sequence 15, Appl
11	1741.5	96.1	594	3	US-09-377-155-7	Sequence 7, Appli
12	1741.5	96.1	594	4	US-09-669-974-7	Sequence 7, Appli
13	1728.5	95.3	592	3	US-09-377-155-17	Sequence 17, Appl
14	1728.5	95.3	592	4	US-09-669-974-17	Sequence 17, Appl
15	1728.5	95.3	598	3	US-09-377-155-5	Sequence 5, Appli
16	1728.5	95.3	598	3	US-09-377-155-13	Sequence 13, Appl
17	1728.5	95.3	598	4	US-09-669-974-5	Sequence 5, Appli
18	1728.5	95.3	598	4	US-09-669-974-13	Sequence 13, Appl
19	1715.5	94.6	589	3	US-09-377-155-19	Sequence 19, Appl
20	1715.5	94.6	589	4	US-09-669-974-19	Sequence 19, Appl
21	608	33.5	1098	1	US-08-409-995-2	Sequence 2, Appli
22	608	33.5	1098	3	US-08-685-467-2	Sequence 2, Appli
23	608	33.5	1098	3	US-09-377-155-32	Sequence 32, Appl
24	608	33.5	1098	3	US-08-913-942-2	Sequence 2, Appli
25	608	33.5	1098	4	US-09-669-974-32	Sequence 32, Appl
26	608	33.5	1098	4	US-09-268-347-44	Sequence 44, Appl
27	590	32.5	658	1	US-08-409-995-5	Sequence 5, Appli
28	590	32.5	658	3	US-08-685-467-5	Sequence 5, Appli
29	590	32.5	658	3	US-08-913-942-5	Sequence 5, Appli
30	581.5	32.1	1094	4	US-09-268-347-32	Sequence 32, Appl
31	569	31.4	2353	3	US-09-377-155-33	Sequence 33, Appl
32	569	31.4	2353	3	US-08-913-942-4	Sequence 4, Appli
33	569	31.4	2353	4	US-09-669-974-33	Sequence 33, Appl
34	569	31.4	2411	4	US-09-268-347-36	Sequence 36, Appl
35	568.5	31.4	2354	4	US-09-268-347-47	Sequence 47, Appl
36	549	30.3	1912	1	US-08-409-995-4	Sequence 4, Appli
37	549	30.3	1912	3	US-08-685-467-4	Sequence 4, Appli
38	529.5	29.2	607	1	US-08-409-995-6	Sequence 6, Appli
39	529.5	29.2	607	3	US-08-685-467-6	Sequence 6, Appli
40	529.5	29.2	607	3	US-08-913-942-6	Sequence 6, Appli
41	412.5	22.8	679	3	US-08-913-942-15	Sequence 15, Appl
42	412.5	22.8	679	4	US-09-268-347-26	Sequence 26, Appl
43	339	18.7	1004	4	US-09-268-347-30	Sequence 30, Appl
44	331.5	18.3	1002	4	US-09-268-347-24	Sequence 24, Appl
45	325	17.9	1104	4	US-09-268-347-28	Sequence 28, Appl

ALIGNMENTS

RESULT 1
US-09-377-155-11
; Sequence 11, Application US/09377155
; Patent No. 6197312
; GENERAL INFORMATION:
; APPLICANT: PEAK, Ian Richard Anselm
; APPLICANT: JENNINGS, Michael Paul
; APPLICANT: MOXON, E. Richard
; TITLE OF INVENTION: NOVEL SURFACE ANTIGEN
; FILE REFERENCE: 065064/0128
; CURRENT APPLICATION NUMBER: US/09/377,155
; CURRENT FILING DATE: 1999-08-19
; PRIOR APPLICATION NUMBER: PCT/AU98/01031
; PRIOR FILING DATE: 1998-12-14
; PRIOR APPLICATION NUMBER: GB 9726398.2
; PRIOR FILING DATE: 1997-12-12
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 591
; TYPE: PRT
; ORGANISM: Neisseria meningitidis
US-09-377-155-11

Query Match 99.7%; Score 1808; DB 3; Length 591;
Best Local Similarity 99.7%; Pred. No. 3.8e-141;
Matches 355; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 NNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKDKG 60
              :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        236 DNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKDKG 295

Qy         61 ENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        296 ENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKG 355

Qy        121 TTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGK 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        356 TTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGK 415

Qy        181 MDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGDALNVGSKKD 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        416 MDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGDALNVGSKKD 475

Qy        241 NKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAYL 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        476 NKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAYL 535

Qy        301 PGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 356
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        536 PGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 591


RESULT 2
US-09-377-155-21
; Sequence 21, Application US/09377155
; Patent No. 6197312
; GENERAL INFORMATION:
;  APPLICANT:  PEAK, Ian Richard Anselm
;  APPLICANT:  JENNINGS, Michael Paul
;  APPLICANT:  MOXON, E. Richard
;  TITLE OF INVENTION:  NOVEL SURFACE ANTIGEN
;  FILE REFERENCE:  065064/0128
;  CURRENT APPLICATION NUMBER: US/09/377,155
;  CURRENT FILING DATE:  1999-08-19
;  PRIOR APPLICATION NUMBER:  PCT/AU98/01031
;  PRIOR FILING DATE:  1998-12-14
;  PRIOR APPLICATION NUMBER:  GB 9726398.2
;  PRIOR FILING DATE:  1997-12-12
;  NUMBER OF SEQ ID NOS:  33
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 21
;   LENGTH: 591
;   TYPE: PRT
;   ORGANISM: Neisseria meningitidis
US-09-377-155-21

  Query Match             99.7%;  Score 1808;  DB 3;  Length 591;
  Best Local Similarity   99.7%;  Pred. No. 3.8e-141;
  Matches  355;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 NNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKDKG 60
              :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        236 DNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKDKG 295

Qy         61 ENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        296 ENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKG 355

Qy        121 TTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGK 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        356 TTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGK 415

Qy        181 MDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGDALNVGSKKD 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        416 MDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGDALNVGSKKD 475

Qy        241 NKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAYL 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        476 NKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAYL 535

Qy        301 PGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 356
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        536 PGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 591


RESULT 3
US-09-669-974-11
; Sequence 11, Application US/09669974
; Patent No. 6333173
; GENERAL INFORMATION:
;  APPLICANT: PEAK, Ian Richard Anselm
;  APPLICANT:  JENNINGS, Michael Paul
;  APPLICANT:  MOXON, E. Richard
;  TITLE OF INVENTION: NOVEL SURFACE ANTIGEN
;  FILE REFERENCE: 065064/0128
;  CURRENT APPLICATION NUMBER: US/09/669,974
;  CURRENT FILING DATE:  2000-09-26
;  PRIOR APPLICATION NUMBER: US 09/377,155
;  PRIOR FILING DATE: 1999-08-19
;  PRIOR APPLICATION NUMBER: PCT/AU98/01031
;  PRIOR FILING DATE: 1998-12-14
;  PRIOR APPLICATION NUMBER: GB 9726398.2
;  PRIOR FILING DATE: 1997-12-12
;  NUMBER OF SEQ ID NOS: 33
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
;   LENGTH: 591
;   TYPE: PRT
;   ORGANISM: Neisseria meningitidis
US-09-669-974-11

  Query Match             99.7%;  Score 1808;  DB 4;  Length 591;
  Best Local Similarity   99.7%;  Pred. No. 3.8e-141;
  Matches  355;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 NNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKDKG 60
              :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        236 DNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKDKG 295

Qy         61 ENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        296 ENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKG 355

Qy        121 TTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGK 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        356 TTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGK 415

Qy        181 MDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGDALNVGSKKD 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        416 MDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGDALNVGSKKD 475

Qy        241 NKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAYL 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        476 NKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAYL 535

Qy        301 PGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 356
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        536 PGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 591


RESULT 4
US-09-669-974-21
; Sequence 21, Application US/09669974
; Patent No. 6333173
; GENERAL INFORMATION:
;  APPLICANT: PEAK, Ian Richard Anselm
;  APPLICANT:  JENNINGS, Michael Paul
;  APPLICANT:  MOXON, E. Richard
;  TITLE OF INVENTION: NOVEL SURFACE ANTIGEN
;  FILE REFERENCE: 065064/0128
;  CURRENT APPLICATION NUMBER: US/09/669,974
;  CURRENT FILING DATE:  2000-09-26
;  PRIOR APPLICATION NUMBER: US 09/377,155
;  PRIOR FILING DATE: 1999-08-19
;  PRIOR APPLICATION NUMBER: PCT/AU98/01031
;  PRIOR FILING DATE: 1998-12-14
;  PRIOR APPLICATION NUMBER: GB 9726398.2
;  PRIOR FILING DATE: 1997-12-12
;  NUMBER OF SEQ ID NOS: 33
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 21
;   LENGTH: 591
;   TYPE: PRT
;   ORGANISM: Neisseria meningitidis
US-09-669-974-21

  Query Match             99.7%;  Score 1808;  DB 4;  Length 591;
  Best Local Similarity   99.7%;  Pred. No. 3.8e-141;
  Matches  355;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 NNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKDKG 60
              :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        236 DNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKDKG 295

Qy         61 ENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKG 120
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        296 ENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKG 355

Qy        121 TTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGK 180
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        356 TTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGK 415

Qy        181 MDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGDALNVGSKKD 240
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        416 MDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGDALNVGSKKD 475

Qy        241 NKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAYL 300
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        476 NKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAYL 535

Qy        301 PGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 356
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        536 PGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 591


RESULT 5
US-09-377-155-2
; Sequence 2, Application US/09377155
; Patent No. 6197312
; GENERAL INFORMATION:
;  APPLICANT:  PEAK, Ian Richard Anselm
;  APPLICANT:  JENNINGS, Michael Paul
;  APPLICANT:  MOXON, E. Richard
;  TITLE OF INVENTION:  NOVEL SURFACE ANTIGEN
;  FILE REFERENCE:  065064/0128
;  CURRENT APPLICATION NUMBER: US/09/377,155
;  CURRENT FILING DATE:  1999-08-19
;  PRIOR APPLICATION NUMBER:  PCT/AU98/01031
;  PRIOR FILING DATE:  1998-12-14
;  PRIOR APPLICATION NUMBER:  GB 9726398.2
;  PRIOR FILING DATE:  1997-12-12
;  NUMBER OF SEQ ID NOS:  33
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 592
;   TYPE: PRT
;   ORGANISM: Neisseria meningitidis
US-09-377-155-2

  Query Match             99.5%;  Score 1804;  DB 3;  Length 592;
  Best Local Similarity   99.4%;  Pred. No. 8.1e-141;
  Matches  354;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 NNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKDKG 60
              :|||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        237 DNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGVKTSVIKEKDGKLVTGKDKG 296

Qy         61 ENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        297 ENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKG 356

Qy        121 TTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGK 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        357 TTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGK 416

Qy        181 MDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGDALNVGSKKD 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        417 MDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGDALNVGSKKD 476

Qy        241 NKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAYL 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        477 NKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAYL 536

Qy        301 PGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 356
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        537 PGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 592


RESULT 6
US-09-669-974-2
; Sequence 2, Application US/09669974
; Patent No. 6333173
; GENERAL INFORMATION:
;  APPLICANT: PEAK, Ian Richard Anselm
;  APPLICANT:  JENNINGS, Michael Paul
;  APPLICANT:  MOXON, E. Richard
;  TITLE OF INVENTION: NOVEL SURFACE ANTIGEN
;  FILE REFERENCE: 065064/0128
;  CURRENT APPLICATION NUMBER: US/09/669,974
;  CURRENT FILING DATE:  2000-09-26
;  PRIOR APPLICATION NUMBER: US 09/377,155
;  PRIOR FILING DATE: 1999-08-19
;  PRIOR APPLICATION NUMBER: PCT/AU98/01031
;  PRIOR FILING DATE: 1998-12-14
;  PRIOR APPLICATION NUMBER: GB 9726398.2
;  PRIOR FILING DATE: 1997-12-12
;  NUMBER OF SEQ ID NOS: 33
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 592
;   TYPE: PRT
;   ORGANISM: Neisseria meningitidis
US-09-669-974-2

  Query Match             99.5%;  Score 1804;  DB 4;  Length 592;
  Best Local Similarity   99.4%;  Pred. No. 8.1e-141;
  Matches  354;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 NNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKDKG 60
              :|||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        237 DNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGVKTSVIKEKDGKLVTGKDKG 296

Qy         61 ENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        297 ENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKG 356

Qy        121 TTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGK 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        357 TTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGK 416

Qy        181 MDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGDALNVGSKKD 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        417 MDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGDALNVGSKKD 476

Qy        241 NKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAYL 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        477 NKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAYL 536

Qy        301 PGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 356
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        537 PGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 592


RESULT 7
US-09-377-155-9
; Sequence 9, Application US/09377155
; Patent No. 6197312
; GENERAL INFORMATION:
;  APPLICANT:  PEAK, Ian Richard Anselm
;  APPLICANT:  JENNINGS, Michael Paul
;  APPLICANT:  MOXON, E. Richard
;  TITLE OF INVENTION:  NOVEL SURFACE ANTIGEN
;  FILE REFERENCE:  065064/0128
;  CURRENT APPLICATION NUMBER: US/09/377,155
;  CURRENT FILING DATE:  1999-08-19
;  PRIOR APPLICATION NUMBER:  PCT/AU98/01031
;  PRIOR FILING DATE:  1998-12-14
;  PRIOR APPLICATION NUMBER:  GB 9726398.2
;  PRIOR FILING DATE:  1997-12-12
;  NUMBER OF SEQ ID NOS:  33
;  SOFTWARE:  PatentIn Ver. 2.0
```

; SEQ ID NO 9
; LENGTH: 594
; TYPE: PRT
; ORGANISM: Neisseria meningitidis
US-09-377-155-9

Query Match 96.7%; Score 1753.5; DB 3; Length 594;
Best Local Similarity 97.2%; Pred. No. 1.2e-136;
Matches 347; Conservative 4; Mismatches 5; Indels 1; Gaps 1;

```
Qy          1 NNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKDKG 60
              :|||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||
Db        238 DNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKRTEVKIGAKTSVIKEKDGKLVTGKDKG 297

Qy         61 ENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKG 120
              || ||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        298 ENDSSTDKGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKG 357

Qy        121 TTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGK 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        358 TTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGK 417

Qy        181 MDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGD-ALNVGSKK 239
              |||||||||||||||||||||||||||||||||||||||||||||||| : |||||||
Db        418 MDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGALNVGSKD 477

Qy        240 DNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAY 299
               |||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        478 ANKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNHIDNVDGNARAGIAQAIATAGLVQAY 537

Qy        300 LPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 356
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        538 LPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 594
```

RESULT 8
US-09-669-974-9
; Sequence 9, Application US/09669974
; Patent No. 6333173
; GENERAL INFORMATION:
; APPLICANT: PEAK, Ian Richard Anselm
; APPLICANT: JENNINGS, Michael Paul
; APPLICANT: MOXON, E. Richard
; TITLE OF INVENTION: NOVEL SURFACE ANTIGEN
; FILE REFERENCE: 065064/0128
; CURRENT APPLICATION NUMBER: US/09/669,974
; CURRENT FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US 09/377,155
; PRIOR FILING DATE: 1999-08-19
; PRIOR APPLICATION NUMBER: PCT/AU98/01031
; PRIOR FILING DATE: 1998-12-14
; PRIOR APPLICATION NUMBER: GB 9726398.2
; PRIOR FILING DATE: 1997-12-12
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 594
; TYPE: PRT
; ORGANISM: Neisseria meningitidis
US-09-669-974-9

Query Match 96.7%; Score 1753.5; DB 4; Length 594;
Best Local Similarity 97.2%; Pred. No. 1.2e-136;
Matches 347; Conservative 4; Mismatches 5; Indels 1; Gaps 1;

```
Qy          1 NNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKDKG 60
              :|||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||
Db        238 DNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKRTEVKIGAKTSVIKEKDGKLVTGKDKG 297

Qy         61 ENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKG 120
              || ||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        298 ENDSSTDKGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKG 357

Qy        121 TTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGK 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        358 TTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGK 417

Qy        181 MDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGD-ALNVGSKK 239
              |||||||||||||||||||||||||||||||||||||||||||||||| : |||||||
Db        418 MDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGALNVGSKD 477

Qy        240 DNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAY 299
               |||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        478 ANKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNHIDNVDGNARAGIAQAIATAGLVQAY 537

Qy        300 LPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 356
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        538 LPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 594
```

RESULT 9
US-09-377-155-15
; Sequence 15, Application US/09377155
; Patent No. 6197312
; GENERAL INFORMATION:
; APPLICANT: PEAK, Ian Richard Anselm
; APPLICANT: JENNINGS, Michael Paul
; APPLICANT: MOXON, E. Richard
; TITLE OF INVENTION: NOVEL SURFACE ANTIGEN
; FILE REFERENCE: 065064/0128
; CURRENT APPLICATION NUMBER: US/09/377,155
; CURRENT FILING DATE: 1999-08-19
; PRIOR APPLICATION NUMBER: PCT/AU98/01031
; PRIOR FILING DATE: 1998-12-14
; PRIOR APPLICATION NUMBER: GB 9726398.2
; PRIOR FILING DATE: 1997-12-12
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
; LENGTH: 599
; TYPE: PRT
; ORGANISM: Neisseria meningitidis
US-09-377-155-15

Query Match 96.4%; Score 1748.5; DB 3; Length 599;
Best Local Similarity 97.2%; Pred. No. 3.1e-136;
Matches 347; Conservative 2; Mismatches 7; Indels 1; Gaps 1;

```
Qy          1 NNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKDKG 60
              :||||| |||||||||||||||||||||||:|||||||||||||||||||||| ||
Db        243 DNVDFVHTYDTVEFLSADTKTTTVNVESKDNGKRTEVKIGAKTSVIKEKDGKLVTGKGKG 302

Qy         61 ENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 ENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKG 362

Qy        121 TTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGK 180
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        363 TTATVSKDDQGNITVKYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGK 422

Qy        181 MDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDG-DALNVGSKK 239
              |||||||||||||||||||||||||||||||||||||||||||||||   |||||||
Db        423 MDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDKGALNVGSKD 482

Qy        240 DNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAY 299
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        483 ANKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAY 542

Qy        300 LPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 356
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        543 LPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 599
```

RESULT 10
US-09-669-974-15
; Sequence 15, Application US/09669974

```
; Patent No. 6333173
; GENERAL INFORMATION:
;  APPLICANT: PEAK, Ian Richard Anselm
;  APPLICANT:  JENNINGS, Michael Paul
;  APPLICANT:  MOXON, E. Richard
;  TITLE OF INVENTION: NOVEL SURFACE ANTIGEN
;  FILE REFERENCE: 065064/0128
;  CURRENT APPLICATION NUMBER: US/09/669,974
;  CURRENT FILING DATE:  2000-09-26
;  PRIOR APPLICATION NUMBER: US 09/377,155
;  PRIOR FILING DATE: 1999-08-19
;  PRIOR APPLICATION NUMBER: PCT/AU98/01031
;  PRIOR FILING DATE: 1998-12-14
;  PRIOR APPLICATION NUMBER: GB 9726398.2
;  PRIOR FILING DATE: 1997-12-12
;  NUMBER OF SEQ ID NOS: 33
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
;   LENGTH: 599
;   TYPE: PRT
;   ORGANISM: Neisseria meningitidis
US-09-669-974-15

  Query Match             96.4%;  Score 1748.5;  DB 4;  Length 599;
  Best Local Similarity   97.2%;  Pred. No. 3.1e-136;
  Matches  347;  Conservative    2;  Mismatches    7;  Indels    1;  Gaps    1;

Qy          1 NNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKDKG 60
              :|||||  |||||||||||||||||||||||||:||||||||||||||||||||||| ||
Db        243 DNVDFVHTYDTVEFLSADTKTTTVNVESKDNGKRTEVKIGAKTSVIKEKDGKLVTGKGKG 302

Qy         61 ENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 ENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKG 362

Qy        121 TTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGK 180
              ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        363 TTATVSKDDQGNITVKYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGK 422

Qy        181 MDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDG-DALNVGSKK 239
              |||||||||||||||||||||||||||||||||||||||||||||||||    |||||||
Db        423 MDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDKGALNVGSKD 482

Qy        240 DNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAY 299
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        483 ANKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAY 542

Qy        300 LPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 356
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        543 LPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 599


RESULT 11
US-09-377-155-7
; Sequence 7, Application US/09377155
; Patent No. 6197312
; GENERAL INFORMATION:
;  APPLICANT:  PEAK, Ian Richard Anselm
;  APPLICANT:  JENNINGS, Michael Paul
;  APPLICANT:  MOXON, E. Richard
;  TITLE OF INVENTION:  NOVEL SURFACE ANTIGEN
;  FILE REFERENCE:  065064/0128
;  CURRENT APPLICATION NUMBER: US/09/377,155
;  CURRENT FILING DATE:  1999-08-19
;  PRIOR APPLICATION NUMBER:  PCT/AU98/01031
;  PRIOR FILING DATE:  1998-12-14
;  PRIOR APPLICATION NUMBER:  GB 9726398.2
;  PRIOR FILING DATE:  1997-12-12
;  NUMBER OF SEQ ID NOS:  33
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 7
;   LENGTH: 594
;   TYPE: PRT
;   ORGANISM: Neisseria meningitidis
US-09-377-155-7

  Query Match             96.1%;  Score 1741.5;  DB 3;  Length 594;
  Best Local Similarity   96.9%;  Pred. No. 1.2e-135;
  Matches  346;  Conservative    2;  Mismatches    8;  Indels    1;  Gaps    1;

Qy          1 NNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKDKG 60
              :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db        238 DNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKGKD 297

Qy         61 ENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        298 ENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKG 357

Qy        121 TTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGK 180
              ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        358 TTATVSKDDQGNITVKYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGK 417

Qy        181 MDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGD-ALNVGSKK 239
              |||||||||||||||||||||||||| |||||||||||||||||||| : |||||||
Db        418 MDETVNINAGNNIEITRNGKNIDIATSMAPQFSSVSLGAGADAPTLSVDDEGALNVGSKD 477

Qy        240 DNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAY 299
              ||||||||||| |||||||||||||||||||||||||||||||||||||||||||||||
Db        478 TNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAY 537

Qy        300 LPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 356
              |||||||||| |||||||||||||||||||||||||||||||||||||||||||||||
Db        538 LPGKSMMAIGGDTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 594


RESULT 12
US-09-669-974-7
; Sequence 7, Application US/09669974
; Patent No. 6333173
; GENERAL INFORMATION:
;  APPLICANT: PEAK, Ian Richard Anselm
;  APPLICANT:  JENNINGS, Michael Paul
;  APPLICANT:  MOXON, E. Richard
;  TITLE OF INVENTION: NOVEL SURFACE ANTIGEN
;  FILE REFERENCE: 065064/0128
;  CURRENT APPLICATION NUMBER: US/09/669,974
;  CURRENT FILING DATE:  2000-09-26
;  PRIOR APPLICATION NUMBER: US 09/377,155
;  PRIOR FILING DATE: 1999-08-19
;  PRIOR APPLICATION NUMBER: PCT/AU98/01031
;  PRIOR FILING DATE: 1998-12-14
;  PRIOR APPLICATION NUMBER: GB 9726398.2
;  PRIOR FILING DATE: 1997-12-12
;  NUMBER OF SEQ ID NOS: 33
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
;   LENGTH: 594
;   TYPE: PRT
;   ORGANISM: Neisseria meningitidis
US-09-669-974-7

  Query Match             96.1%;  Score 1741.5;  DB 4;  Length 594;
  Best Local Similarity   96.9%;  Pred. No. 1.2e-135;
  Matches  346;  Conservative    2;  Mismatches    8;  Indels    1;  Gaps    1;

Qy          1 NNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKDKG 60
              :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db        238 DNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKGKD 297

Qy         61 ENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        298 ENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKG 357

Qy        121 TTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGK 180
              ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        358 TTATVSKDDQGNITVKYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGK 417
```

```
Qy        181 MDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGD-ALNVGSKK 239
              |||||||||||||||||||||||||||| |||||||||||||||||||| : |||||||
Db        418 MDETVNINAGNNIEITRNGKNIDIATSMAPQFSSVSLGAGADAPTLSVDDEGALNVGSKD 477

Qy        240 DNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAY 299
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        478 TNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAY 537

Qy        300 LPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 356
              |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        538 LPGKSMMAIGGDTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 594


RESULT 13
US-09-377-155-17
; Sequence 17, Application US/09377155
; Patent No. 6197312
; GENERAL INFORMATION:
;  APPLICANT:  PEAK, Ian Richard Anselm
;  APPLICANT:  JENNINGS, Michael Paul
;  APPLICANT:  MOXON, E. Richard
;  TITLE OF INVENTION:  NOVEL SURFACE ANTIGEN
;  FILE REFERENCE:  065064/0128
;  CURRENT APPLICATION NUMBER: US/09/377,155
;  CURRENT FILING DATE:  1999-08-19
;  PRIOR APPLICATION NUMBER:  PCT/AU98/01031
;  PRIOR FILING DATE:  1998-12-14
;  PRIOR APPLICATION NUMBER:  GB 9726398.2
;  PRIOR FILING DATE:  1997-12-12
;  NUMBER OF SEQ ID NOS:  33
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 17
;   LENGTH: 592
;   TYPE: PRT
;   ORGANISM: Neisseria meningitidis
US-09-377-155-17

  Query Match             95.3%;  Score 1728.5;  DB 3;  Length 592;
  Best Local Similarity   96.4%;  Pred. No. 1.4e-134;
  Matches  344;  Conservative    3;  Mismatches    9;  Indels    1;  Gaps    1;

Qy          1 NNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKDKG 60
              :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db        236 DNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKGKG 295

Qy         61 ENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKG 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||| |
Db        296 ENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTKVTFASGNG 355

Qy        121 TTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGK 180
              ||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        356 TTATVSKDDQGNITVKYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGK 415

Qy        181 MDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGD-ALNVGSKK 239
              |||||||||||||||||||||||||||||||||||||||||||||||||| : |||||||
Db        416 MDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGALNVGSKD 475

Qy        240 DNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAY 299
              ||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||||
Db        476 ANKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVNGNARAGIAQAIATAGLVQAY 535

Qy        300 LPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 356
              |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        536 LPGKSMMAIGGGTYLGEAGYAIGYSSISAGGNWIIKGTASGNSRGHFGASASVGYQW 592


RESULT 14
US-09-669-974-17
; Sequence 17, Application US/09669974
; Patent No. 6333173
; GENERAL INFORMATION:
;  APPLICANT: PEAK, Ian Richard Anselm
;  APPLICANT:  JENNINGS, Michael Paul
;  APPLICANT:  MOXON, E. Richard
;  TITLE OF INVENTION: NOVEL SURFACE ANTIGEN
;  FILE REFERENCE: 065064/0128
;  CURRENT APPLICATION NUMBER: US/09/669,974
;  CURRENT FILING DATE:  2000-09-26
;  PRIOR APPLICATION NUMBER: US 09/377,155
;  PRIOR FILING DATE: 1999-08-19
;  PRIOR APPLICATION NUMBER: PCT/AU98/01031
;  PRIOR FILING DATE: 1998-12-14
;  PRIOR APPLICATION NUMBER: GB 9726398.2
;  PRIOR FILING DATE: 1997-12-12
;  NUMBER OF SEQ ID NOS: 33
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17
;   LENGTH: 592
;   TYPE: PRT
;   ORGANISM: Neisseria meningitidis
US-09-669-974-17

  Query Match             95.3%;  Score 1728.5;  DB 4;  Length 592;
  Best Local Similarity   96.4%;  Pred. No. 1.4e-134;
  Matches  344;  Conservative    3;  Mismatches    9;  Indels    1;  Gaps    1;

Qy          1 NNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKDKG 60
              :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db        236 DNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKGKG 295

Qy         61 ENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKG 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||| |
Db        296 ENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTKVTFASGNG 355

Qy        121 TTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGK 180
              ||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        356 TTATVSKDDQGNITVKYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGK 415

Qy        181 MDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGD-ALNVGSKK 239
              |||||||||||||||||||||||||||||||||||||||||||||||||| : |||||||
Db        416 MDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGALNVGSKD 475

Qy        240 DNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAY 299
              ||||||||||||||||||||||||||||||||||||||||:||||||||||||||||||||
Db        476 ANKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVNGNARAGIAQAIATAGLVQAY 535

Qy        300 LPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 356
              |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        536 LPGKSMMAIGGGTYLGEAGYAIGYSSISAGGNWIIKGTASGNSRGHFGASASVGYQW 592


RESULT 15
US-09-377-155-5
; Sequence 5, Application US/09377155
; Patent No. 6197312
; GENERAL INFORMATION:
;  APPLICANT:  PEAK, Ian Richard Anselm
;  APPLICANT:  JENNINGS, Michael Paul
;  APPLICANT:  MOXON, E. Richard
;  TITLE OF INVENTION:  NOVEL SURFACE ANTIGEN
;  FILE REFERENCE:  065064/0128
;  CURRENT APPLICATION NUMBER: US/09/377,155
;  CURRENT FILING DATE:  1999-08-19
;  PRIOR APPLICATION NUMBER:  PCT/AU98/01031
;  PRIOR FILING DATE:  1998-12-14
;  PRIOR APPLICATION NUMBER:  GB 9726398.2
;  PRIOR FILING DATE:  1997-12-12
;  NUMBER OF SEQ ID NOS:  33
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 5
;   LENGTH: 598
;   TYPE: PRT
;   ORGANISM: Neisseria meningitidis
US-09-377-155-5
```

Query Match 95.3%; Score 1728.5; DB 3; Length 598;
Best Local Similarity 95.8%; Pred. No. 1.4e-134;
Matches 342; Conservative 4; Mismatches 10; Indels 1; Gaps 1;

```
Qy          1 NNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKDKG 60
              :|||||||||||||||||||||||||||||||:|||||||||||||||||||||| ||
Db        242 DNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKRTEVKIGAKTSVIKEKDGKLVTGKGKG 301

Qy         61 ENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKG 120
              |||||||||||||||||||||||||||||||||||||||||||||||||| |||||| |
Db        302 ENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTKVTFASGNG 361

Qy        121 TTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGK 180
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        362 TTATVSKDDQGNITVKYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGK 421

Qy        181 MDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGD-ALNVGSKK 239
              |||||||||||||||||||||||||||||||||||||||||||||||||| : |||||||
Db        422 MDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGALNVGSKD 481

Qy        240 DNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAY 299
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db        482 ANKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLAQAY 541

Qy        300 LPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 356
              |||||||||||||||||||||||||| |||:||||||||||||| ||||||||
Db        542 LPGKSMMAIGGGTYRGEAGYAIGYSSISDTGNWVIKGTASGNSRGHFGTSASVGYQW 598
```

Search completed: October 6, 2003, 09:35:59
Job time : 10.5202 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: October 6, 2003, 09:13:35 ; Search time 10.1906 Seconds
(without alignments)
3359.577 Million cell updates/sec

Title: US-09-771-382-37
Perfect score: 1813
Sequence: 1 NNVDFVRTYDTVEFLSADTK.....TASGNSRGHFGASASVGYQW 356

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_76:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1808	99.7	591	2	G81133	adhesin NMB0992 [i
2	1752.5	96.7	592	2	A81888	probable surface f
3	332.5	18.3	2059	2	D82671	surface protein XF
4	327	18.0	1190	2	A82615	surface protein XF
5	325	17.9	1588	2	A86036	probable adhesin Z
6	325	17.9	1588	2	H91188	probable adhesin E
7	314	17.3	1107	2	AC0976	probable autotrans
8	307.5	17.0	658	2	AH0110	probable surface p
9	183.5	10.1	365	2	AB3486	cell surface prote
10	176.5	9.7	1004	2	C82672	surface-exposed ou
11	167.5	9.2	585	2	F90961	flagellin [importe
12	167.5	9.2	585	2	F85809	hypothetical prote
13	163	9.0	298	2	I64138	adhesin homolog HI
14	156.5	8.6	933	2	S41539	fibrinogen-binding
15	156	8.6	584	2	C48658	flagellin - Escher
16	155.5	8.6	989	2	D89852	fibrinogen-binding
17	154	8.5	364	2	AE0169	probable exported
18	153	8.4	4919	2	T31105	hypothetical prote
19	152.5	8.4	364	2	A81019	adhesin/invasin, p
20	152	8.4	1477	2	B43855	high-molecular-wei
21	151.5	8.4	2020	2	C48399	ABC-type transport
22	151	8.3	980	2	H90681	probable flagellin
23	151	8.3	980	2	D85532	probable structura
24	151	8.3	1536	2	A43855	high-molecular-wei
25	148	8.2	1910	2	AF0394	probable adhesin h
26	147	8.1	1335	2	T17508	glycoprotein Vp260
27	146	8.1	1461	2	E90696	hypothetical prote
28	145	8.0	595	2	A48658	flagellin - Escher
29	144.5	8.0	936	2	I40711	sapB protein - Cam
30	144.5	8.0	1361	2	T03415	S-layer protein -
31	144	7.9	1635	2	AI0452	hemolysin [importe
32	144	7.9	2554	2	AB3528	extracellular seri
33	143.5	7.9	906	2	C85739	hypothetical prote
34	143.5	7.9	1011	2	F90879	probable BigA-like
35	143.5	7.9	1052	2	AF2959	conserved hypothet
36	143.5	7.9	1341	2	H98323	hypothetical prote
37	143	7.9	1127	2	T25804	hypothetical prote
38	143	7.9	1519	2	S41525	major ring-forming
39	142	7.8	920	2	I40614	surface array prot
40	142	7.8	1461	2	A85547	hypothetical prote
41	142	7.8	5188	2	B85547	probable RTX famil
42	142	7.8	5291	2	F90696	hypothetical prote
43	141.5	7.8	338	2	D90697	adhesin/invasin-li
44	141.5	7.8	338	2	G85547	adhesin/invasin-li
45	141.5	7.8	1035	2	AD3203	autotransporter pr

ALIGNMENTS

RESULT 1
G81133
adhesin NMB0992 [imported] - Neisseria meningitidis (strain MC58 serogroup B)
C;Species: Neisseria meningitidis
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001
C;Accession: G81133
R;Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B. ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Masignani, V.; Pizza, M. Science 287, 1809-1815, 2000
A;Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.;
A;Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A;Reference number: A81000; MUID:20175755; PMID:10710307
A;Accession: G81133
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-591 <TET>
A;Cross-references: GB:AE002450; GB:AE002098; NID:g7226229; PIDN:AAF41395.1; PID:g722
A;Experimental source: serogroup B, strain MC58
C;Genetics:
A;Gene: NMB0992

Query Match 99.7%; Score 1808; DB 2; Length 591;
Best Local Similarity 99.7%; Pred. No. 4.7e-97;
Matches 355; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 NNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKDKG 60
              :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        236 DNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKDKG 295

Qy         61 ENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        296 ENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKG 355

Qy        121 TTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGK 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        356 TTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGK 415

Qy        181 MDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGDALNVGSKKD 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        416 MDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGDALNVGSKKD 475

Qy        241 NKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAYL 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        476 NKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAYL 535

Qy        301 PGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 356
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        536 PGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 591
```

RESULT 2

```
A81888
probable surface fibril protein NMA1200 [imported] - Neisseria meningitidis (strain Z249
C;Species: Neisseria meningitidis
C;Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001
C;Accession: A81888
R;Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel
; Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,
Nature 404, 502-506, 2000
A;Title: Complete DNA sequence of a serogroup A strain of Neisseria menigitidis Z2491.
A;Reference number: A81775; MUID:20222556; PMID:10761919
A;Accession: A81888
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-592 <PAR>
A;Cross-references: GB:AL162755; GB:AL157959; NID:g7379742; PIDN:CAB84461.1; PID:g737989
A;Experimental source: serogroup A, strain Z2491
C;Genetics:
A;Gene: NMA1200

  Query Match             96.7%;  Score 1752.5;  DB 2;  Length 592;
  Best Local Similarity   97.5%;  Pred. No. 7.5e-94;
  Matches  347;  Conservative    3;  Mismatches    5;  Indels    1;  Gaps    1;

Qy          2 NVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKDKGE 61
              |||||||||||||||||||||||||||||:|||||||||||||||||||||| |||
Db        237 NVDFVRTYDTVEFLSADTKTTTVNVESKDNGKRTEVKIGAKTSVIKEKDGKLVTGKGKGE 296

Qy         62 NGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKGT 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        297 NGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKGT 356

Qy        122 TATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKM 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        357 TATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKM 416

Qy        182 DETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGD-ALNVGSKKD 240
              |||||||||||||:|||||||||||| |||||||||||||||||||| : |||||||
Db        417 DETVNINAGNNIEISRNGKNIDIATSMAPQFSSVSLGAGADAPTLSVDDEGALNVGSKDA 476

Qy        241 NKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAYL 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        477 NKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAYL 536

Qy        301 PGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 356
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        537 PGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 592


RESULT 3
D82671
surface protein XF1529 [imported] - Xylella fastidiosa (strain 9a5c)
C;Species: Xylella fastidiosa
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C;Accession: D82671
R;anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000
A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A;Reference number: A82515; MUID:20365717; PMID:10910347
A;Note: for a complete list of authors see reference number A59328 below
A;Accession: D82671
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-2059 <SIM>
A;Cross-references: GB:AE003982; GB:AE003849; NID:g9106554; PIDN:AAF84338.1; GSPDB:GN001
A;Experimental source: strain 9a5c
R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H
as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigr
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E
A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
, F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri,
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawa
A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silv
M.; Tsuhako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.
A;Reference number: A59328
A;Contents: annotation
C;Genetics:
A;Gene: XF1529

  Query Match             18.3%;  Score 332.5;  DB 2;  Length 2059;
  Best Local Similarity   29.3%;  Pred. No. 2.5e-11;
  Matches   98;  Conservative   53;  Mismatches  112;  Indels   71;  Gaps   12;

Qy         28 SKDNGKKTEVKIGAKTSVIKEKDGKLVTG-KDKGENGSSTDEGEGLVTAKEVIDAVNKAG 86
              :|| |:  |   |  : |     | :  |   |||  | ::    :  ||| :||||
Db       1792 AKDGGRGAESYTGKYSGVQNNTVGTVSVGDAAKGETRSISN----VADAKEAMDAVN--- 1844

Qy         87 WRMKTTTANGQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALN 146
              |    |    | |                          : :  ||  :: :  ||  :
Db       1845 LRQLDAVAQKSNLQTDDM----------------------RHEINNIEDVFKITKGD--S 1880

Qy        147 VNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIAT 206
              : ::  | |    |:|  :   :||  | : ||     |::| | : |  ||
Db       1881 ASSVKGMGVN----AMAIGTNAAVSGTESVALGK---NTNVSADNAVAI-GNG------- 1925

Qy        207 SMTPQFSSVSLGAGADAPTLSVDGDALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKG 266
              |:  : :|||:|:|               ||::     ::|||| |  : |  ||:||
Db       1926 SVADRANSVSVGSG--------------GSER-----QVTNVAAGTADTDAVNVSQLNQ 1965

Qy        267 VAQNLNNRIDNVDGNAR----AGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIG 322
                 | : || |      |:| ||||| | |||: |: | ::|  :|:|::  |:|
Db       1966 GLITAKQYTDGMVGNLRRETSGGVAAAIATANLPQAYVQGRGMTSVGVSSYQGQSAIAVG 2025

Qy        323 YSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 356
              |::|: |:|: | : | |:| | |  | |||||
Db       2026 VSAVSESGHWVFKFSGSANTRSHVGVGAGVGYQW 2059


RESULT 4
A82615
surface protein XF1981 [imported] - Xylella fastidiosa (strain 9a5c)
C;Species: Xylella fastidiosa
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C;Accession: A82615
R;anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Seq
Nature 406, 151-157, 2000
A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A;Reference number: A82515; MUID:20365717; PMID:10910347
A;Note: for a complete list of authors see reference number A59328 below
A;Accession: A82615
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1190 <SIM>
A;Cross-references: GB:AE004017; GB:AE003849; NID:g9107083; PIDN:AAF84783.1; GSPDB:GN
A;Experimental source: strain 9a5c
R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer
as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Fr
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; La
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins
A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.
, F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri,
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawa
A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silv
M.; Tsuhako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.
A;Reference number: A59328
A;Contents: annotation
C;Genetics:
A;Gene: XF1981

  Query Match             18.0%;  Score 327;  DB 2;  Length 1190;
```

```
  Best Local Similarity   27.3%;  Pred. No. 2.6e-11;
  Matches  108;  Conservative   52;  Mismatches  143;  Indels   92;  Gaps   14;

Qy          2 NVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKDKGE 61
              :| :  |||      |:   | :    :  | :|   |    :|    |
Db        848 SVHYYSTYD------GGTQGGNYNGDGATGTRSIAVGVGTLASA----EGATAVGSGAAA 897

Qy         62 NG-SSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKG 120
              :|  ||  |   | : :   |:   | :     |   ||:    :  | ||     : ||
Db        898 SGKGSTAIGRNAVASADGSVALGD-GAKDGARGAESYTGKYSGLQNNTVGTVSVGDASKG 956

Qy        121 TTATVSKDDQGNITVMYDVNVGDALNVNQL----QNSGWNLDSKAVAGSSGKVISGNVSP 176
               | |||     |:    :    ||:|: ||    |::   :|:|  : | |:
Db        957 ETRTVS-----NVADAKEAT--DAVNLRQLDRVAQDANRYVDNKIESLSEGQTF------ 1003

Qy        177 SKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSV-----DGD 231
                 :|: ||              | ||      : || ||  : |       |
Db       1004 --------VKVNSLNN-------------SATP------IAAGVDATAIGVGATASGAD 1035

Qy        232 ALNVGSK----KDNKPV---------------------RITNVAPGVKEGDVTNVAQLK 265
              :: :|:|     ||                      ::|||| |  : |  ||:||
Db       1036 SIAMGNKASASADNAVAIGNHSVADRANTVSVGSAGSERQVTNVAAGTADTDAVNVSQLN 1095

Qy        266 GVAQNLNNRIDNVDGNAR----AGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAI 321
                   | | |: |      |:| ||||| | |||:||: | ::|  :|||::  |:
Db       1096 QGLITAKQYTDGVVGSLRRDTDGGVAAAIATANLPQAYIPGRGMTSVGVSSYRGQSAIAV 1155

Qy        322 GYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 356
              | ||:|: | |: | : | |:|   |  | |||||
Db       1156 GVSSVSESGRWVFKFSGSANTRSQVGIGAGVGYQW 1190


RESULT 5
A86036
probable adhesin Z5029 [imported] - Escherichia coli (strain O157:H7, substrain EDL933)
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C;Accession: A86036
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: A86036
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1588 <STO>
A;Cross-references: GB:AE005174; NID:g12518349; PIDN:AAG58749.1; GSPDB:GN00145; UWGP:Z5
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: Z5029

  Query Match             17.9%;  Score 325;  DB 2;  Length 1588;
  Best Local Similarity   26.7%;  Pred. No. 4.9e-11;
  Matches  115;  Conservative   58;  Mismatches  165;  Indels   92;  Gaps   15;

Qy          1 NNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEK----------D 50
              | :  | |  |  | :  |:  :: | :        : :|||| |  :|          |
Db       1177 NAIGAVATTPTKYFHANSTEEDSLAVGT------DSLAMGAKTIVNGDKGIGIGYGAYVD 1230

Qy         51 GKLVTGKDKGE-------------NGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQ 97
                : |   |               |||:|  |          :||   :       :|:||
Db       1231 ANALNGIAIGSNAQVIHVNSIAIGNGSTTTRGAQTNYTAYNMDAPQNSVGEFSVGSADGQ 1290

Qy         98 -------TGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVM-------------Y 137
                    | ||      |   ||      | | | ||:: | :|| :
Db       1291 RQITNVAAGSAD-----TDAVNV--GQLKVTDAQVSQNTQ-SITNLDNRVTNLDSRVTNI 1342

Qy        138 DVNVGDAL----------NVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKM---DET 184
              :  :|| :          | : :  |    || |:   |      :|:   | :   : |
Db       1343 ENGIGDIVTTGSTKYFKTNTDGVDASAQGKDSVAIGSGSIAAADNSVALGTGSVATEENT 1402

Qy        185 VNINAGNNIEITRN---GKN----IDIATSMTPQFSSVSLGAGADAPTLSVDGDALNVGS 237
              :::  :  |      |  |||    :::|   : :   |      ||    |:|   : :|
Db       1403 ISVGSSTNQRRITNVAAGKNATDAVNVAQLKSSEAGGVRYDTKADG---SIDYSNITLGG 1459

Qy        238 KKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQ-----------NLNNRIDNVDGNARAGI 286
                :    ||:||: ||   || | ||||    |          ::|::   :     ||
Db       1460 -GNGGTTRISNVSAGVNNNDVVNYAQLKQSVQETKQYTDQRMVEMDNKLSKTESKLSGGI 1518

Qy        287 AQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHF 346
              | |:|  || ||| || || :|||||| ||:  |:| | :|  | |: |   | ||:| :
Db       1519 ASAMAMTGLPQAYTPGASMASIGGGTYNGESAVALGVSMVSANGRWVYKLQGSTNSQGEY 1578

Qy        347 GASASVGYQW 356
               |:   | ||
Db       1579 SAALGAGIQW 1588


RESULT 6
H91188
probable adhesin ECs4480 [similarity] - Escherichia coli (strain O157:H7, substrain R
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 02-Nov-2001
C;Accession: H91188
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and g
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: H91188
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1588 <HAY>
A;Cross-references: GB:BA000007; PIDN:BAB37903.1; PID:g13363955; GSPDB:GN00154
A;Experimental source: strain O157:H7, substrain RIMD 0509952
C;Genetics:
A;Gene: ECs4480

  Query Match             17.9%;  Score 325;  DB 2;  Length 1588;
  Best Local Similarity   26.7%;  Pred. No. 4.9e-11;
  Matches  115;  Conservative   58;  Mismatches  165;  Indels   92;  Gaps   15;

Qy          1 NNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEK----------D 50
              | :  | |  |  | :  |:  :: | :        : :|||| |  :|          |
Db       1177 NAIGAVATTPTKYFHANSTEEDSLAVGT------DSLAMGAKTIVNGDKGIGIGYGAYVD 1230

Qy         51 GKLVTGKDKGE-------------NGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQ 97
                : |   |               |||:|  |          :||   :       :|:||
Db       1231 ANALNGIAIGSNAQVIHVNSIAIGNGSTTTRGAQTNYTAYNMDAPQNSVGEFSVGSADGQ 1290

Qy         98 -------TGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVM-------------Y 137
                    | ||      |   ||      | | | ||:: | :|| :
Db       1291 RQITNVAAGSAD-----TDAVNV--GQLKVTDAQVSQNTQ-SITNLDNRVTNLDSRVTNI 1342

Qy        138 DVNVGDAL----------NVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKM---DET 184
              :  :|| :          | : :  |    || |:   |      :|:   | :   : |
Db       1343 ENGIGDIVTTGSTKYFKTNTDGVDASAQGKDSVAIGSGSIAAADNSVALGTGSVATEENT 1402

Qy        185 VNINAGNNIEITRN---GKN----IDIATSMTPQFSSVSLGAGADAPTLSVDGDALNVGS 237
              :::  :  |      |  |||    :::|   : :   |      ||    |:|   : :|
Db       1403 ISVGSSTNQRRITNVAAGKNATDAVNVAQLKSSEAGGVRYDTKADG---SIDYSNITLGG 1459

Qy        238 KKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQ-----------NLNNRIDNVDGNARAGI 286
                :    ||:||: ||   || | ||||    |          ::|::   :     ||
Db       1460 -GNGGTTRISNVSAGVNNNDVVNYAQLKQSVQETKQYTDQRMVEMDNKLSKTESKLSGGI 1518

Qy        287 AQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHF 346
              | |:|  || ||| || || :|||||| ||:  |:| | :|  | |: |   | ||:| :
Db       1519 ASAMAMTGLPQAYTPGASMASIGGGTYNGESAVALGVSMVSANGRWVYKLQGSTNSQGEY 1578

Qy        347 GASASVGYQW 356
               |:   | ||
Db       1579 SAALGAGIQW 1588
```

```
RESULT 7
AC0976
probable autotransporter sapB [imported] - Salmonella enterica subsp. enterica serovar T
C;Species: Salmonella enterica subsp. enterica serovar Typhi
A;Note: this species has also been called Salmonella typhi
C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C;Accession: AC0976
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A;Reference number: AB0502; MUID:21534947; PMID:11677608
A;Accession: AC0976
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1107 <PAR>
A;Cross-references: GB:AL513382; PIDN:CAD03303.1; PID:g16504923; GSPDB:GN00176
C;Genetics:
A;Gene: sapB

  Query Match             17.3%;  Score 314;  DB 2;  Length 1107;
  Best Local Similarity   26.8%;  Pred. No. 1.4e-10;
  Matches  109;  Conservative   55;  Mismatches  165;  Indels   78;  Gaps   14;

Qy         16 SADTKTTTVNVESKDNGKKTEVK------IGAKTSVIKEKDGKLVTGKDKGE-------- 61
              |  :  :  |  :|   | || |       ||  | |: :    :  | :
Db        713 STEEDSLAVGTDSLAMGAKTIVNADAGIGIGLNTLVMADAINGIAIGSNARANHANSIAM 772

Qy         62 -NGSSTDEG-EGLVTAKEVIDAVNKAGWRMKTTTANGQ-------TGQADKFETVTSGTN 112
               ||| |  | :   ||  :    |  |       : :||        | ||     |   |
Db        773 GNGSQTTRGAQTDYTAYNMDTPQNSVG-EFSVGSEDGQRQITNVAAGSAD-----TDAVN 826

Qy        113 VTFASGKGTTATVSKDDQGNITVMYDVN------------VGDALNVNQLQNSGWNLD-- 158
              |     | | | ||:: |    :   |:           :|| :     :    | |
Db        827 V--GQLKVTDAQVSRNTQSITNLNTQVSNLDTRVTNIENGIGDIVTTGSTKYFKTNTDGA 884

Qy        159 ------SKAVAGSSGKVISGNVSPSKGK---MDETVNINAGNNIEITR---------NGK 200
                   : :||  || : :   | : |     ||   :: |:: :  |          |
Db        885 DANAQGADSVAIGSGSIAAAENSVALGTNSVADEANTVSVGSSTQQRRITNVAAGVNNTD 944

Qy        201 NIDIATSMTPQFSSVSLGAGADAPTLSVDGDALNVGSKKDNKPVRITNVAPGVKEGDVTN 260
               :::|     :  ||     ||    ||:   ||:|        || ||:  | : |  |
Db        945 AVNVAQLKASEAGSVRYETNADG---SVNYSVLNLGDGSGG-TTRIGNVSAAVNDTDAVN 1000

Qy        261 VAQLKGVAQ-----------NLNNRIDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIG 309
              ||||   :           :|::|  ::     ||| |:| ||| ||| || :| :|
Db       1001 YAQLKRSVEEANTYTDQKMGEMNSKIKGIENKMSGGIASAMAMAGLPQAYAPGANMTSIA 1060

Qy        310 GGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 356
              |||: ||:  ||| | :|: | |: |   : ||:| : |:   |:||
Db       1061 GGTFNGESAVAIGVSMVSESGGWVYKLQGTSNSQGDYSAAIGAGFQW 1107


RESULT 8
AH0110
probable surface protein (partial) YPO0902 [imported] - Yersinia pestis (strain CO92)
C;Species: Yersinia pestis
C;Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 02-Nov-2001
C;Accession: AH0110
R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,
Nature 413, 523-527, 2001
A;Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A;Reference number: AB0001; MUID:21470413; PMID:11586360
A;Accession: AH0110
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-658 <KUR>
A;Cross-references: GB:AL590842; PIDN:CAC89747.1; PID:g15978974; GSPDB:GN00175
C;Genetics:
A;Gene: YPO0902

  Query Match             17.0%;  Score 307.5;  DB 2;  Length 658;
  Best Local Similarity   28.4%;  Pred. No. 1.7e-10;
  Matches   97;  Conservative   48;  Mismatches  121;  Indels   75;  Gaps   12;

Qy         40 GAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTG 99
              ||  |     || :  |:  |:||   :   |                    :  :|  :|
Db        369 GAAASA----DGSVAIGQGSGDNGRGVENYIG----------------KYSNASNTSSG 407

Qy        100 QADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDS 159
                    ||: |   |     | | |||    |       :   ||:|: ||
Db        408 ------TVSVGNTAT-----GETRTVSNVADG-------LQATDAVNLRQLDG------- 442

Qy        160 KAVAGSSGKVISGNVSPSKGKMDETVNINAGNNI-EITRNGKNIDI--ATSMTPQFSSVS 216
                  :|  |:  ||| :   |    :|  : : : :  | |     | |:    :| :
Db        443 ---IAASIVVVENNVSGLQNGTDGMFQVNNSSGLAKPSATGANSATGGAGSVASGNNSTA 499

Qy        217 LGAGADAP----------TLSVDGDALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKG 266
               |:|| |           :::   ::::|||  : :  :||||||  :  |  |  |||
Db        500 FGSGAKATAANSAALGANSVADRANSVSVGSVGNER--QITNVAPATQGTDAVNFDQLKS 557

Qy        267 VAQN----LNNRIDNVDGNAR-------AGIAQAIATAGLVQAYLPGKSMMAIGGGTYRG 315
              ::       | |   :  : |       ||||  |:: | | | |  | ||  ||  :|||
Db        558 ISNQTNAYTNQRYSELKQDLRKQNSVLSAGIASAMSMASLTQPYTSGSSMTTIGAASYRG 617

Qy        316 EAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 356
              ::  ::| ||||| | |: |  || |::| ||   |||||
Db        618 QSALSLGVSSISDSGRWVSKLQASSNTQGDFGIGVGVGYQW 658


RESULT 9
AB3486
cell surface protein [imported] - Brucella melitensis (strain 16M)
C;Species: Brucella melitensis
C;Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 01-Feb-2002
C;Accession: AB3486
R;DelVecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanov
.; Mazur, M.; Goltsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Let
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A;Title: The genome sequence of the facultative intracellular pathogen Brucella melit
A;Reference number: AD3252; PMID:11756688
A;Accession: AB3486
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-365 <KUR>
A;Cross-references: GB:AE008917; PIDN:AAL53053.1; PID:g17983913; GSPDB:GN00190
A;Experimental source: strain 16M
C;Genetics:
A;Gene: BMEI1872
A;Map position: I

  Query Match             10.1%;  Score 183.5;  DB 2;  Length 365;
  Best Local Similarity   23.1%;  Pred. No. 0.0012;
  Matches   77;  Conservative   38;  Mismatches   82;  Indels  137;  Gaps   11;

Qy        128 DDQGNIT-VMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVN 186
              ::  ||  :::|:  |  :      : |   ||:|:  :|  | | :|:  :|      :
Db         29 ENLANINDILHDIESGGGIKYFHANSIG--ADSRALGTNSIAVGSDSVASGEG------S 80

Qy        187 INAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGA----------------------DAP 224
              |: ||  : : :|      :  |  :||:||||:                      |||
Db         81 ISVGNGAQASAHGSVALGENAAAPDANSVALGAGSKTSEVVATKGTTINGQYYDFAGDAP 140

Qy        225 --------------------------TLSVDGDALNV---------------------- 235
                                       | :|:|  ||
Db        141 SGTVSVGDKGAERTITNVAAGRISVESTDAVNGSQLNAVNQAIENLAAGVTENDKFSVKY 200

Qy        236 -----GSKKDN---------KPVRITNVAPGVKEGDVTNVAQLK-GVAQNLNNRIDNVDG 280
                  |:||::          || : ||| || : |  ||:||| |::  |       |
Db        201 DRHSDGTKKNSMTLQGWDSATPVVLANVADGVHKNDAVNVSQLKAGLSTTLGEAKAYTDQ 260

Qy        281 NARAGIAQA----------------------IATAGLVQAYL-----PGKSMMAIGGG 311
```

```
Db        261 TALQTLDQANAYTDKKFGKLNEDIVATRIEARQAAAIGLAAASLRYDDRPGKISAAIGGG 320

Qy        312 TYRGEAGYAIG-------------YSSISDGGNW 332
Db        321 FWRGEGAVALGLGHTSEDQRMRSNLSAATSGGNW 354
```

RESULT 10
C82672
surface-exposed outer membrane protein XF1516 [imported] - Xylella fastidiosa (strain 9a
C;Species: Xylella fastidiosa
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C;Accession: C82672
R;anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000
A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A;Reference number: A82515; MUID:20365717; PMID:10910347
A;Note: for a complete list of authors see reference number A59328 below
A;Accession: C82672
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1004 <SIM>
A;Cross-references: GB:AE003981; GB:AE003849; NID:g9106543; PIDN:AAF84325.1; GSPDB:GN001
A;Experimental source: strain 9a5c
R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H
as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigr
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E
A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
, F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
M.; Tsuhako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A;Reference number: A59328
A;Contents: annotation
C;Genetics:
A;Gene: XF1516

  Query Match   9.7%; Score 176.5; DB 2; Length 1004;
  Best Local Similarity 25.8%; Pred. No. 0.01;
  Matches   98; Conservative   58; Mismatches 141; Indels   83; Gaps   18;

```
Qy         27 ESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKDKGE-NGSSTDEGEGLV--TAKEVI---- 79
Db        139 QSANNG----VALGSNSTVSGVNSVALGAGSMASELNVISVGGGDGVTGPAVRRIVNVGD 194

Qy         80 -----DAVNKAGWRMKTTTANGQTGQADKFETV-----TSGTNVTFASGKGTTATVS-KD 128
Db        195 GIGNNDAVNKSQLDGVTASVN---DVAASVKTIALTNQVTGSSVASASGKESTAIGSGAQ 251

Qy        129 DQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNIN 188
Db        252 AVADNTVAF----GGRAIANAVGASALGFDSHAKGINSTTVGTQSVSLGQGGVSLGYNSF 307

Qy        189 AG----NNIEITRNG----KNIDIATSMTPQFSSVSLGAGADAPTLSV----DGDALNVG 236
Db        308 VGEGSFNGLALGSNSLVLLQGVD----------SVALGSGSMASEPNVVSVGSGDGLRGP 357

Qy        237 SKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNV---------DGNARAGIA 287
Db        358 AVR-----RIVNVGDGIGNNDAVNKSQLDGVTASVNDVVASVKNIAGAIQITGSGVASVS 412

Qy        288 QAIATAGLVQAYLPGKSMMAIG--------GGTYRGEAGYAIGYSSISDGGNWIIKGTA- 338
Db        413 GQDSTAAGASAQAAGDSSIALGARSRANAIGSSALGVDGHALGANSTALGG----QSTAI 468

Qy        339 --SGNSRGH---FGASASVG 353
Db        469 SEGGTSLGYNSFVGQSATNG 488
```

RESULT 11
F90961
flagellin [imported] - Escherichia coli (strain O157:H7, substrain RIMD 0509952)
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001
C;Accession: F90961
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and g
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: F90961
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-585 <HAY>
A;Cross-references: GB:BA000007; PIDN:BAB36085.1; PID:g13362130; GSPDB:GN00154
A;Experimental source: strain O157:H7, substrain RIMD 0509952
C;Genetics:
A;Gene: ECs2662
C;Superfamily: flagellin

  Query Match   9.2%; Score 167.5; DB 2; Length 585;
  Best Local Similarity 23.0%; Pred. No. 0.018;
  Matches   85; Conservative   54; Mismatches 161; Indels   69; Gaps   16;

```
Qy         10 DTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEG 69
Db        121 DEIDRVSGQTQFNGVNVLAKDGSMK--IQVGANDGETITIDLKKIDSDTLGLNGFNVN-G 177

Qy         70 EGLVTAK-EVIDAVNKAGWRMKTTTANGQTGQAD-KFETVTSGTNVTFAS-GKGTTATVS 126
Db        178 KGTITNKAATVSDLTSAGAKLNTT-----TGLYDLKTENTLLTTDAAFDKLGNGDKVTV- 231

Qy        127 KDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSS---------------GKVIS 171
Db        232 ----GGVDYTYNAKSGDFTTTK--STAGTGVDAAAQAADSASKRDALAATLHADVGKSVN 285

Qy        172 GNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGA----------G 220
Db        286 GSYTTKDGTV--SFETDSAGNITIGGSQAYVDDAGNLTTNNAGSAAKADMKALLKAASEG 343

Qy        221 ADAPTLSVDGDALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDG 280
Db        344 SDGASLTFNGTEYTIAKA---TPATTTPVAPLIPGGITYQATVSKDVVLS--------ET 392

Qy        281 NARAGIAQAIATAGLVQ---AYLPGKSMMAI-------GGGTYRGEAGYAIGYSSISDGG 330
Db        393 KAAAATSSITFNSGVLSKTIGFTAGESSDAAKSYVDDKGGIT--NVADYTVSYSVNKDNG 450

Qy        331 NWIIKGTAS 339
Db        451 SVTVAGYAS 459
```

RESULT 12
F85809
hypothetical protein fliC [imported] - Escherichia coli (strain O157:H7, substrain ED
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C;Accession: F85809
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apoda
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: F85809
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-585 <STO>
A;Cross-references: GB:AE005174; NID:g12516024; PIDN:AAG56938.1; GSPDB:GN00145; UWGP:
A;Experimental source: strain O157:H7, substrain EDL933

```
C;Genetics:
A;Gene: fliC
C;Superfamily: flagellin

  Query Match             9.2%;  Score 167.5;  DB 2;  Length 585;
  Best Local Similarity   23.0%;  Pred. No. 0.018;
  Matches   85;  Conservative   54;  Mismatches  161;  Indels   69;  Gaps   16;

Qy         10 DTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEG 69
              | :: :|  |:   ||| :||   |  :::||       | | :     | || : : |
Db        121 DEIDRVSGQTQFNGVNVLAKDGSMK--IQVGANDGETITIDLKKIDSDTLGLNGFNVN-G 177

Qy         70 EGLVTAK-EVIDAVNKAGWRMKTTTANGQTGQAD-KFETVTSGTNVTFAS-GKGTTATVS 126
              :| :| |   :  :  || :: ||     ||  | | |     |:  |   | |   ||
Db        178 KGTITNKAATVSDLTSAGAKLNTT-----TGLYDLKTENTLLTTDAAFDKLGNGDKVTV- 231

Qy        127 KDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSS---------------GKVIS 171
                  | :   |:   ||         :|  :|: | |  |                || ::
Db        232 ----GGVDYTYNAKSGDFTTTK--STAGTGVDAAAQAADSASKRDALAATLHADVGKSVN 285

Qy        172 GNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGA-----------G 220
              |: :   | :  :   ::  || |  :   :| | ::|   :  :  |           |
Db        286 GSYTTKDGTV--SFETDSAGNITIGGSQAYVDDAGNLTTNNAGSAAKADMKALXKAASEG 343

Qy        221 ADAPTLSVDGDALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDG 280
              :|  :|: :|     :      |  | ||| :  |       | |  :          :
Db        344 SDGASLTFNGTEYTIAKA---TPATTTPVAPLIPGGITYQATVSKDVVLS--------ET 392

Qy        281 NARAGIAQAIATAGLVQ---AYLPGKSMMAI-------GGGTYRGEAGYAIGYSSISDGG 330
              | |  :     :|::     :  |:|  |        || |    | | : ||   | |
Db        393 KAAAATSSITFNSGVLSKTIGFTAGESSDAAKSYVDDKGGIT--NVADYTVSYSVNKDNG 450

Qy        331 NWIIKGTAS 339
              :  : | ||
Db        451 SVTVAGYAS 459


RESULT 13
I64138
adhesin homolog HI1732 - Haemophilus influenzae (strain Rd KW20)
C;Species: Haemophilus influenzae
C;Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 24-Oct-1997
C;Accession: I64138
R;Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A
; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J
, D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M.
Science 269, 496-512, 1995
A;Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,
A;Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
A;Reference number: A64000; MUID:95350630; PMID:7542800
A;Accession: I64138
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-298 <TIGR>
A;Cross-references: GB:U32846; GB:L42023; NID:g1574588; PID:g1574589; TIGR:HI1732

  Query Match             9.0%;  Score 163;  DB 2;  Length 298;
  Best Local Similarity   49.3%;  Pred. No. 0.014;
  Matches   35;  Conservative   11;  Mismatches   23;  Indels    2;  Gaps    1;

Qy          2 NVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKD--K 59
              ||| |  || |||:: |  |  | : :|:||| ||||   ||||||: :|||:|||
Db        221 NVDLVAGYDNVEFITGDKNTLDVVLTAKENGKTTEVKFTPKTSVIKDNNGKLLTGKQLKD 280

Qy         60 GENGSSTDEGE 70
                 |::|:  |
Db        281 ANTGTATNATE 291


RESULT 14
S41539
fibrinogen-binding protein - Staphylococcus aureus
N;Alternate names: clumping factor
C;Species: Staphylococcus aureus
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 15-Oct-1999
C;Accession: S41539; S36630
R;McDevitt, D.; Francois, P.; Vaudaux, P.; Foster, T.J.
Mol. Microbiol. 11, 237-248, 1994
A;Title: Molecular characterization of the clumping factor (fibrinogen receptor) of S
A;Reference number: S41539; MUID:94224142; PMID:8170386
A;Accession: S41539
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-933 <MCD>
A;Cross-references: EMBL:Z18852; NID:g397525; PIDN:CAA79304.1; PID:g397526

  Query Match             8.6%;  Score 156.5;  DB 2;  Length 933;
  Best Local Similarity   21.0%;  Pred. No. 0.13;
  Matches   83;  Conservative   55;  Mismatches  164;  Indels   93;  Gaps   13;

Qy         24 VNVESKDNGKKTEVKIGAKTSVI----------KEKDG--KLVTGKDKGENGSSTDEGE 70
              :|:: |:      :  ||  : ::          || |     ||  |  | | :::
Db          1 MNMKKKEKHAIRKKSIGVASVLVGTLIGFGLLSSKEADASENSVTQSDSASNESKSNDSS 60

Qy         71 GLVTAKEVIDAVNKAGWRMKTTTANGQT----GQADKFETVTSGTNVTF----ASGKGTT 122
                ||     |  | :  :  : | ||:|      | :  | :| || |      :|: ||
Db         61 S-VSAAPKTDDTNVSDTKTSSNTNNGETSVAQNPAQQETTQSSSTNATTEETPVTGEATT 119

Qy        123 ATVSK-----------------DDQGNITVMYDVNVGDALNVNQLQNSGWNL------D 158
                | ::                   :  | |   | |   ::|  |   :  |:
Db        120 TTTNQANTPATTQSSNTNAEELVNQTSNETTFNDTNTVSSVNSPQNSTNAENVSTTQDTS 179

Qy        159 SKAVAGSSGKVISGNVSPSKGKMDETVNIN-----------------AGNNI--EITRN 198
              ::|   ::        : :|  ::: || :                  || :|  ::|
Db        180 TEATPSNNESAPQSTDASNKDVVNQAVNTSAPRMRAFSLAAVAADAPAAGTDITNQLTNV 239

Qy        199 GKNIDIATSMTP-QFSSVSLGAGADAPTLSVDGDALNVGSKKDNKPVRITNVA--PGVKE 255
                ||  |:: | |   | |  |   |  :| ||   :   |:     :|: |  | :
Db        240 TVGIDSGTTVYPHQAGYVKLNYGFSVPNSAVKGDTFKITVPKELNLNGVTSTAKVPPIMA 299

Qy        256 GDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAYL-------------PG 302
              ||           | | | : : |||        : |   |:| |              |
Db        300 GD----------QVLANGVIDSDGNVIYTFTDYVNTKDDVKATLTMPAYIDPENVKKTG 348

Qy        303 KSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGT 337
              :| | |:        : |       |  ||||
Db        349 NVTLATGIGSTTANKTVLVDYEKYGKFYNLSIKGT 383


RESULT 15
C48658
flagellin - Escherichia coli (strain U5-41)
C;Species: Escherichia coli
C;Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 26-Aug-1999
C;Accession: C48658
R;Schoenhals, G.; Whitfield, C.
J. Bacteriol. 175, 5395-5402, 1993
A;Title: Comparative analysis of flagellin sequences from Escherichia coli strains po
A;Reference number: A48658; MUID:93374833; PMID:8366026
A;Accession: C48658
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-584 <SCH>
A;Cross-references: GB:L07388; NID:g290438; PIDN:AAA23798.1; PID:g290439
C;Superfamily: flagellin

  Query Match             8.6%;  Score 156;  DB 2;  Length 584;
  Best Local Similarity   23.8%;  Pred. No. 0.081;
  Matches   88;  Conservative   52;  Mismatches  158;  Indels   72;  Gaps   18;

Qy         10 DTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEG 69
              | :: :|  |:   ||| :||   |  :::||       | | :     | || : : |
Db        121 DEIDRVSGQTQFNGVNVLAKDGSMK--IQVGANDGETITIDLKKIDSDTLGLNGFNVN-G 177

Qy         70 EGLVTAK-EVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFAS-GKGTTATVSK 127
              :| :| |   :  :  || :: |||  |  |   | |     |:  |   | |   ||
```

```
Db        178 KGTITNKAATVSDLTSAGAKLNTTT--GLYGL--KTENTLLTTDAAFDKLGNGDKVTVGG 233

Qy        128 DDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSS---------------GKVISG 172
              |       :|:   ||          :|  :|: | |  |                || ::|
Db        234 VD------IYNAKSGDFTTTK--STAGTGVDAAAQATDSAKKRDALAATLHADVGKSVNG 285

Qy        173 NVSPSKGKMDETVNINAGNNIEITRNGKN--IDIATSMTPQFSSVSLGA----------- 219
              : :   | :   |  :|||   ||  |    :| | ::|   :  :  |
Db        286 SYTTKDGTV-SFVTDSAGN---ITIGGSQAYVDDAGNLTTNNAGSARKADMKALLKAASE 341

Qy        220 GADAPTLSVDGDALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVD 279
              |:|  :|: :|    :       |   : ||| :  |        | |  :        :
Db        342 GSDGASLTFNGTEYTIAKA---TPATTSPVAPLIPGGITYQATVSKDVVLS--------E 390

Qy        280 GNARAGIAQAIATAGLVQ---AYLPGKSMMAI-------GGGTYRGEAGYAIGYSSISDG 329
                | |  :     :|::     :  |:|  |        || |    | | : ||   |
Db        391 TKAAAATSSVTFNSGVLSKTIGFTAGESSDAAKSYVDDKGGIT--NVADYTVSYSVNKDN 448

Qy        330 GNWIIKGTAS 339
              |:  : | ||
Db        449 GSVTVAGYAS 458


Search completed: October  6, 2003, 09:33:39
Job time : 12.1906 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: October 6, 2003, 09:06:20 ; Search time 5.49755 Seconds
(without alignments)
3045.266 Million cell updates/sec

Title: US-09-771-382-37
Perfect score: 1813
Sequence: 1 NNVDFVRTYDTVEFLSADTK.....TASGNSRGHFGASASVGYQW 356

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	151.5	8.4	2003	1	YDBA_ECOLI	P33666 escherichia
2	141.5	7.8	1953	1	BIGA_SALTY	P25927 salmonella
3	140	7.7	1655	1	OMPB_RICCN	Q9kka3 r outer mem
4	136.5	7.5	550	1	FLIC_SHIFL	Q08860 shigella fl
5	136.5	7.5	930	1	PMP8_CHLPN	Q9z393 chlamydia p
6	136.5	7.5	1300	1	120K_RICRI	P14914 rickettsia
7	136.5	7.5	1654	1	OMPB_RICRI	Q53047 r outer mem
8	135.5	7.5	497	1	FLIC_ECOLI	P04949 escherichia
9	134.5	7.4	1286	1	AIDA_ECOLI	Q03155 escherichia
10	132.5	7.3	1608	1	HLYA_SERMA	P15320 serratia ma
11	131	7.2	434	1	YADA_YERPS	P10858 yersinia ps
12	131	7.2	1398	1	TOP2_PLAFK	P41001 plasmodium
13	130.5	7.2	1569	1	YPJA_ECOLI	P52143 escherichia
14	130	7.2	1577	1	HLYA_PROMI	P16466 proteus mir
15	130	7.2	2249	1	OMPA_RICRI	P15921 rickettsia
16	129.5	7.1	716	1	CX80_EUPOC	Q9n9x3 euplotes oc
17	129	7.1	3178	1	YS89_CAEEL	Q09624 caenorhabdi
18	128.5	7.1	1567	1	ICEN_XANCT	P18127 xanthomonas
19	128	7.1	1645	1	OMPB_RICTY	P96989 r outer mem
20	126.5	7.0	575	1	FLA2_CAMJE	P22251 campylobact
21	126.5	7.0	933	1	SLAP_CAMFE	P35827 campylobact
22	126	6.9	928	1	HXA2_HAEIN	P45354 haemophilus
23	125.5	6.9	713	1	CDG2_PAEMA	P31835 paenibacill
24	125.5	6.9	760	1	AMY_CLOAB	P23671 clostridium
25	125.5	6.9	917	1	HXA3_HAEIN	P45355 haemophilus
26	125.5	6.9	928	1	PM10_CHLPN	Q9rb65 chlamydia p
27	125	6.9	1656	1	OMPB_RICJA	O06653 r outer mem
28	124.5	6.9	455	1	YADA_YEREN	P31489 yersinia en
29	124	6.8	1025	1	SLAP_CAUCR	P35828 caulobacter
30	123.5	6.8	2660	1	YEEJ_ECO57	Q8x8v7 escherichia
31	123	6.8	581	1	FIBP_ADE05	P11818 human adeno
32	123	6.8	860	1	VG12_BPB03	Q37893 bacteriopha
33	123	6.8	1325	1	YDEK_ECOLI	P32051 escherichia
34	122.5	6.8	1007	1	Y741_CHLMU	Q9pjt6 chlamydia m
35	122.5	6.8	1153	1	PVDB_PLAKN	P50493 plasmodium
36	122.5	6.8	1861	1	APU_THETU	P38536 t amylopull
37	122.5	6.8	2021	1	OMPA_RICCN	Q52657 rickettsia
38	122	6.7	762	1	SLAP_ACEKI	P22258 acetogenium
39	121.5	6.7	367	1	FLC2_PROMI	P42273 proteus mir
40	121	6.7	670	1	GR78_YARLI	Q99170 yarrowia li
41	121	6.7	918	1	YMJB_CAEEL	P34487 caenorhabdi
42	121	6.7	1694	1	IGA0_HAEIN	P44969 haemophilus
43	121	6.7	1702	1	IGA2_HAEIN	P45384 haemophilus
44	121	6.7	2334	1	WAPA_BACSU	Q07833 bacillus su
45	120.5	6.6	575	1	FLB2_CAMJE	P22252 campylobact

ALIGNMENTS

RESULT 1
YDBA_ECOLI
ID YDBA_ECOLI STANDARD; PRT; 2003 AA.
AC P33666; P76087; P76088; P76856; P76857; P76859;
DT 01-FEB-1994 (Rel. 28, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein ydbA.
GN YDBA OR B1401/B1405.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=97251357; PubMed=9097039;
RA Aiba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K., Itoh T.,
RA Kasai H., Kashimoto K., Kimura S., Kitakawa M., Kitagawa M.,
RA Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,
RA Nakade S., Nakamura Y., Nashimoto H., Nishio Y., Oshima T., Saito N.,
RA Sampei G., Seki Y., Sivasundaram S., Tagami H., Takeda J.,
RA Takemoto K., Takeuchi Y., Wada C., Yamamoto Y., Horiuchi T.;
RT "A 570-kb DNA sequence of the Escherichia coli K-12 genome
RT corresponding to the 28.0-40.1 min region on the linkage map.";
RL DNA Res. 3:363-377(1996).
RN [3]
RP SEQUENCE OF 464-2003 FROM N.A.
RC STRAIN=K12;
RX MEDLINE=92190338; PubMed=1665988;
RA Moszer I., Glaser P., Danchin A.;
RT "Multiple IS insertion sequences near the replication terminus in'
RT Escherichia coli K-12.";
RL Biochimie 73:1361-1374(1991).
CC -!- SIMILARITY: TO S.TYPHIMURIUM ORF NEAR CYSG (AC P25928).
CC -!- CAUTION: THIS IS A CONCEPTUAL TRANSLATION; THE GENE CODING FOR
CC THIS PROTEIN IS INTERRUPTED BY A HYBRID IS2D/IS30C ELEMENT
CC BETWEEN AMINO ACIDS 839 AND 840.
CC ---

```
CC   ----------------------------------------------------------------------
DR   EMBL; AE000237; AAC74483.1; ALT_SEQ.
DR   EMBL; AE000237; AAC74487.1; ALT_SEQ.
DR   EMBL; D90778; BAA15009.1; ALT_SEQ.
DR   EMBL; D90778; BAA18880.1; ALT_SEQ.
DR   EMBL; D90779; BAA18881.1; ALT_SEQ.
DR   EMBL; X62680; -; NOT_ANNOTATED_CDS.
DR   EcoGene; EG11307; ydbA.
KW   Hypothetical protein; Complete proteome.
FT   CONFLICT    489    489       I -> V (IN REF. 2).
FT   CONFLICT    495    495       I -> V (IN REF. 2).
SQ   SEQUENCE   2003 AA;  205949 MW;  B83A12C8B53220EE CRC64;

  Query Match             8.4%;  Score 151.5;  DB 1;  Length 2003;
  Best Local Similarity   23.4%;  Pred. No. 0.36;
  Matches  101;  Conservative   54;  Mismatches  154;  Indels  123;  Gaps   23;

Qy          3 VDFVRTYDTVEFLSADTKT------------TTVNVESKDNGKKTEVKIGAKTSV---IK 47
              :|       :| :|||  |             |  |    :||| |   :  | |     |
Db        189 IDEANNTVALEGVSADGATKWQYNHNGELVITGDNATVNNNGKTT---VDGKDSTGTEIN 245

Qy         48 EKDGKLVTGKDKGENG----------SSTDEGEGLVTAKE------VIDA----VNKAGW 87
               :||::   |   :|          |:| : :| :|  :        ||     ||  |
Db        246 GNNGKVIQDGDLDVSGGGHGIDITGDSATVDNKGTMTVTDPESMGIQIDGDKAIVNNEG- 304

Qy         88 RMKTTTANGQTGQADKFETVT-------------SGTNVTFASGK-------------- 119
                ::|  || ||     :  |              :|| :   :||
Db        305 --ESTITNGGTGTQINGDDATANNNGKTTVDGKDSTGTEINGNNGKVIQDGDLDVSGGGH 362

Qy        120 -----GTTATVSKDDQGNITVMYDVNVGDALNVNQ--LQNSGWNLDSKAVAGSSGKVISG 172
                  | :|||  |::| :||    ::|  :: :|  : | |   :|    | :|  |:|
Db        363 GIDITGDSATV--DNKGTMTVTDPESIGIQVDGDQAVVNNEG---ESAITNGGTGTQING 417

Qy        173 N--VSPSKGK-----MDETVNINAGNN--------IEITRNGKNIDI-ATSMTPQFSSVS 216
              :   : : ||      | |    ||||        :::: |  |||   | |
Db        418 DDATANNNGKTTVDGKDSTGTEIAGNNGKVIQDGDLDVSGGGHGIDITGDSATVDNKGTM 477

Qy        217 LGAGADAPTLSVDGDALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNR-I 275
                ::  : :|||   | :: ::    |||   |          |: |     ||
Db        478 TVTDPESIGIQIDGDQAIVNNEGES---TITNGGTG---------TQINGNDATANNSGK 525

Qy        276 DNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIK 335
              |||    |  :     |:|   | | |:   ||         | |  :| | |    |
Db        526 TTVDGKDSTG-TKIAGNIGIVN--LDG-SLTVTGG---------AHGVENIGDNGTVNNK 572

Qy        336 GTASGNSRGHFG 347
              |    :  |  |
Db        573 GDIVVSDTGSIG 584


RESULT 2
BIGA_SALTY
ID   BIGA_SALTY     STANDARD;      PRT;  1953 AA.
AC   P25927; P25928; Q9XCQ3;
DT   01-MAY-1992 (Rel. 22, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Putative surface-exposed virulence protein bigA precursor.
GN   BIGA OR STM3478.
OS   Salmonella typhimurium.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Salmonella.
OX   NCBI_TaxID=602;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 14028;
RA   Stojiljkovic I., Valentine P., Heffron F.;
RT   "Salmonella typhimurium rhs homolog.";
RL   Submitted (MAR-1999) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX   MEDLINE=21534948; PubMed=11677609;
RA   McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA   Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA   Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA   Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA   Waterston R., Wilson R.K.;
RT   "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT   LT2.";
RL   Nature 413:852-856(2001).
RN   [3]
RP   SEQUENCE OF 1-765 FROM N.A.
RC   STRAIN=LT2;
RX   MEDLINE=91100301; PubMed=1987123;
RA   Wu J.Y., Siegel L.M., Kredich N.M.;
RT   "High-level expression of Escherichia coli NADPH-sulfite reductase:
RT   requirement for a cloned cysG plasmid to overcome limiting siroheme
RT   cofactor.";
RL   J. Bacteriol. 173:325-333(1991).
CC   -!- CAUTION: Ref.3 sequence differs from that shown due to frameshifts
CC       in positions 414 and 732.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AF133696; AAD39458.1; -.
DR   EMBL; AE008859; AAL22340.1; -.
DR   EMBL; M64606; AAA27042.1; ALT_FRAME.
DR   EMBL; M64606; AAA27043.1; ALT_FRAME.
DR   StyGene; SG10437; bigA.
KW   Virulence; Repeat; Signal; Complete proteome.
FT   SIGNAL        1     27       POTENTIAL.
FT   CHAIN        28   1953       PUTATIVE SURFACE-EXPOSED VIRULENCE
FT                                PROTEIN BIGA.
FT   DOMAIN      101    252       15 X 11 AA TANDEM REPEATS.
FT   REPEAT      101    103       1 (INCOMPLETE).
FT   REPEAT      104    113       2 (INCOMPLETE).
FT   REPEAT      114    122       3 (INCOMPLETE).
FT   REPEAT      123    133       4.
FT   REPEAT      134    144       5.
FT   REPEAT      145    155       6.
FT   REPEAT      156    166       7.
FT   REPEAT      167    177       8.
FT   REPEAT      178    188       9.
FT   REPEAT      189    199       10.
FT   REPEAT      200    210       11.
FT   REPEAT      211    221       12.
FT   REPEAT      222    232       13.
FT   REPEAT      233    243       14.
FT   REPEAT      244    252       15 (INCOMPLETE).
FT   CONFLICT    207    207       D -> DRGDDDVTPPDD (IN REF. 1).
FT   CONFLICT    514    514       A -> R (IN REF. 3).
FT   CONFLICT   1698   1698       D -> N (IN REF. 1).
FT   CONFLICT   1795   1798       QYLE -> ITLQ (IN REF. 1).
FT   CONFLICT   1836   1837       SA -> T (IN REF. 1).
SQ   SEQUENCE   1953 AA;  200150 MW;  611B3F1C954D91AE CRC64;

  Query Match             7.8%;  Score 141.5;  DB 1;  Length 1953;
  Best Local Similarity   23.6%;  Pred. No. 1.3;
  Matches   86;  Conservative   45;  Mismatches  133;  Indels  101;  Gaps   18;

Qy          9 YDTVEFL-SADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKDKGEN----- 62
              ||:  :|  ||  || :: : : :     : |       ::: | :::| |:  |
Db        333 YDSRGYLIIADDNTTVISGDDQAHNSDRGMDISG-----QDRTGVIISG-DRTVNTLTGD 386

Qy         63 GSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKGTT 122
                | ||   |:|                   : :| |       || : |     || |||
Db        387 SSVTDGATGMV-----------------ISGDGTTNTISGHSTVDNATG-ALISGNGTT 427

Qy        123 ATVSKD---DQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKG 179
```

```
Db        428 TNFAGDIAVSGGGTAIIIDGDNATIKNTGTSDISG--------AGSTGTVIDGNNARVNN 479

Qy        180 KMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGD---ALNVG 236

Db        480 DGDMTIT-DGGTGGHITGDNVVIDNAGSTTV--------SGADATALYIEGDNALVINEG 530

Qy        237 SKKDNKPVRITNVAPGVK-EGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGL 295

Db        531 NQ------TISGGAVGTRIDGDDAHTTNTGDIA---------VDGAGSAAVIINGDNGSL 575

Qy        296 VQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIK-GTASGNSRGHFGASA---- 350

Db        576 TQA----------------------GDLLVTDGAMGIITYGT--GNEAKNTGNATVRDA 610

Qy        351 -SVGY 354

Db        611 DSVGF 615


RESULT 3
OMPB_RICCN
ID   OMPB_RICCN     STANDARD;      PRT;  1655 AA.
AC   Q9KKA3; Q9KK98; Q9XC45;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Outer membrane protein B precursor (168 kDa surface-layer protein)
DE   (Surface protein antigen) (Cell surface antigen 5) (Sca5) (rOmpB)
DE   (rOmp B) [Contains: 120 kDa surface-exposed protein (Surface protein
DE   antigen) (120 kDa outer membrane protein ompB); 32 kDa beta peptide].
GN   OMPB OR RC1085.
OS   Rickettsia conorii.
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC   Rickettsiaceae; Rickettsieae; Rickettsia.
OX   NCBI_TaxID=781;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Malish 7;
RX   MEDLINE=21442074; PubMed=11557893;
RA   Ogata H., Audic S., Renesto-Audiffren P., Fournier P.-E., Barbe V.,
RA   Samson D., Roux V., Cossart P., Weissenbach J., Claverie J.-M.,
RA   Raoult D.;
RT   "Mechanisms of evolution in Rickettsia conorii and R. prowazekii.";
RL   Science 293:2093-2098(2001).
RN   [2]
RP   SEQUENCE OF 33-1649 FROM N.A.
RC   STRAIN=Indian tick typhus, and Malish 7;
RX   MEDLINE=20393643; PubMed=10939649;
RA   Roux V., Raoult D.;
RT   "Phylogenetic analysis of members of the genus Rickettsia using the
RT   gene coding the outer-membrane protein rOmpB (ompB).";
RL   Int. J. Syst. Evol. Microbiol. 50:1449-1455(2000).
RN   [3]
RP   SEQUENCE OF 353-1655 FROM N.A.
RC   STRAIN=Malish 7;
RA   Stenos J., Walker D.;
RT   "The rickettsial outer membrane protein A and B genes of Rickettsia
RT   australis, the most divergent rickettsia of the spotted fever group.";
RL   Submitted (MAY-1999) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: THE 120 KDA SURFACE-EXPOSED PROTEIN IS A MAJOR
CC       STRUCTURAL PROTEIN WHICH MAY PLAY A ROLE AS A RICKETTSIAL
CC       VIRULENCE FACTOR AND/OR IMMUNOGEN DURING INFECTION (BY
CC       SIMILARITY).
CC   -!- FUNCTION: THE 32 KDA BETA PEPTIDE MAY SERVE AS A MEMBRANE ANCHOR
CC       (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: CELL WALL. THIS BACTERIUM IS COVERED BY A S-
CC       LAYER WITH HEXAGONAL SYMMETRY (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE RICKETTSIAE OMPA/OMPB FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE008659; AAL03623.1; -.
DR   EMBL; AF123721; AAF34124.1; -.
DR   EMBL; AF123726; AAF34129.1; -.
DR   EMBL; AF149110; AAD39533.1; -.
DR   PIR; E97835; E97835.
DR   InterPro; IPR006315; Autotransport.
DR   InterPro; IPR005546; Autotransporter.
DR   Pfam; PF03797; Autotransporter; 1.
DR   TIGRFAMs; TIGR01414; autotrans_barl; 2.
KW   Antigen; S-layer; Cell wall; Complete proteome.
FT   CHAIN         1   1334       120 kDa SURFACE-EXPOSED PROTEIN.
FT   CHAIN      1335   1655       32 kDa BETA PEPTIDE.
FT   VARIANT      61     61       P -> A (IN STRAIN INDIAN TICK TYPHUS).
FT   VARIANT      75     75       G -> S (IN STRAIN INDIAN TICK TYPHUS).
FT   VARIANT      78     78       K -> N (IN STRAIN INDIAN TICK TYPHUS).
FT   VARIANT     251    251       V -> A (IN STRAIN INDIAN TICK TYPHUS).
FT   VARIANT     413    413       N -> D (IN STRAIN INDIAN TICK TYPHUS).
FT   VARIANT     959    959       I -> V (IN STRAIN INDIAN TICK TYPHUS).
FT   VARIANT     988    988       A -> T (IN STRAIN INDIAN TICK TYPHUS).
FT   VARIANT    1139   1139       R -> L (IN STRAIN INDIAN TICK TYPHUS).
FT   CONFLICT    353    354       KD -> GH (IN REF. 3).
FT   CONFLICT    776    776       F -> S (IN REF. 3).
FT   CONFLICT   1159   1159       E -> D (IN REF. 3).
FT   CONFLICT   1177   1177       G -> S (IN REF. 3).
FT   CONFLICT   1492   1492       H -> R (IN REF. 3).
SQ   SEQUENCE   1655 AA;  168342 MW;  E49E19377D5FCE37 CRC64;

  Query Match             7.7%;  Score 140;  DB 1;  Length 1655;
  Best Local Similarity   24.3%;  Pred. No. 1.3;
  Matches   98;  Conservative   48;  Mismatches  176;  Indels   82;  Gaps   20;

Qy          2 NVDFVRTYD---TVEFLSADTKTTTVN----VESKDNG-------KKTEVKIGAKTSVIK 47

Db        430 NTAGVITFDANGTLASASADANVAVTNNITAIEASGAGVVQLSGTHAAELRLGNAGSVFK 489

Qy         48 EKDGKLVTGKDKG--------ENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTG 99

Db        490 LADGTVINGKVNQTALVGGALAAGTITLDGSATITG----DIGNAGG----AAALQGITL 541

Qy        100 QADKFETVT---------SGTNVTFASGKGTTATVSKDDQGNITVMYDV-------NVGD 143

Db        542 ANDATKTLTLGGANIIGANGGTINFQANGGTIKLTS--TQNNIVVDFDLAIATDQTGVVD 599

Qy        144 ALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKM-DETVNIN---AGNN--IEITR 197

Db        600 ASSLTNAQTLTINGKIGTVGANNKTLGQFNIGSSKTVLSDGDVAINELVIGNNGAVQFAH 659

Qy        198 NGKNIDIATSMTPQ----FSSV-----SLGAGADAPTLSVDGDALNVGSKKDNKPVRITN 248

Db        660 NTYLITRTTNAAGQGKIIFNPVVNNNTTLATGTNLGSATNPLAEINFGSKGAANVDTVLN 719

Qy        249 VAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDG-NARAGIAQAIATAGLVQ-------AYL 300

Db        720 VGKGVNL-YATNITTTDA---NVGSFIFNAGGTNIVSG------TVGGQQGNKFNTVALD 769

Qy        301 PGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRG 344

Db        770 NGTTVKFLGNATFNGNTTIAAN-STLQIGGNYTADFVASADGTG 812


RESULT 4
FLIC_SHIFL
ID   FLIC_SHIFL     STANDARD;      PRT;   550 AA.
AC   Q08860;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Flagellin.
```

```
GN   FLIC OR SF1966.
OS   Shigella flexneri.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Shigella.
OX   NCBI_TaxID=623;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=IID642;
RX   MEDLINE=94335647; PubMed=8057852;
RA   Tominaga A., Mahmoud M.A.-H., Mukaihara T., Enomoto M.;
RT   "Molecular characterization of intact, but cryptic, flagellin genes
RT   in the genus Shigella.";
RL   Mol. Microbiol. 12:277-285(1994).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=301 / Serotype 2a;
RX   MEDLINE=22272406; PubMed=12384590;
RA   Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
RA   Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
RA   Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
RA   Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA   Yu J.;
RT   "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
RT   through comparison with genomes of Escherichia coli K12 and O157.";
RL   Nucleic Acids Res. 30:4432-4441(2002).
CC   -!- FUNCTION: FLAGELLIN IS THE SUBUNIT PROTEIN WHICH POLYMERIZES TO
CC       FORM THE FILAMENTS OF BACTERIAL FLAGELLA.
CC   -!- SIMILARITY: BELONGS TO THE BACTERIAL FLAGELLIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D16819; BAA04093.1; -.
DR   EMBL; AE015215; AAN43516.1; -.
DR   PIR; S44980; S44980.
DR   InterPro; IPR001029; Flagellin_C.
DR   InterPro; IPR001492; FlagellinN.
DR   Pfam; PF00700; Flagellin_C; 1.
DR   Pfam; PF00669; Flagellin_N; 1.
DR   PRINTS; PR00207; FLAGELLIN.
KW   Flagella.
SQ   SEQUENCE   550 AA;  56636 MW;  CC921C9A8EF200B6 CRC64;

  Query Match             7.5%;  Score 136.5;  DB 1;  Length 550;
  Best Local Similarity   21.8%;  Pred. No. 0.59;
  Matches   73;  Conservative   50;  Mismatches  161;  Indels   51;  Gaps   13;

Qy         10 DTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEG 69
              | :: :|  |:   ||| :||   |  :::||        | | :     | || : : |
Db        121 DEIDRVSGQTQFNGVNVLAKDGSMK--IQVGANDGQTITIDLKKIDSDTLGLNGFNVNGG 178

Qy         70 EGL---VTAKEVIDAVNKAGWRMKTTTANGQTG--QADKFETVTSGTNV--TFASGKGTT 122
                :     :|  : | |       | | : |      :|    |: |  |  |  :| ||
Db        179 GAVANTAASKADLVAANATVVGNKYTVSAGYDAAKASDLLAGVSDGDTVQATINNGFGTA 238

Qy        123 ATVS--KDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGK---VISGNVSPS 177
              |: :  | |  : :  :|     | :| :    |        | : |    | |:
Db        239 ASATNYKYDSASKSYSFDTTTASAADVQKYLTPG--------VGDTAKGTITIDGSAQDV 290

Qy        178 KGKMDETVNINAGNNI------EITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGD 231
              :   |  :  : |: :       :|:||       |:|    |        | |: : |
Db        291 QISSDGKITASNGDKLYIDTTGRLTKNGS----GASLTEASLSTLAANNTKATTIDIGGT 346

Qy        232 ALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNN--------RIDNVDGNAR 283
              :::  :     |  ||    | |   |   |  | |: : ||          ::   |
Db        347 SISF-TGNSTTPDTITYSVTGAK---VDQAAFDKAVSTSGNNVDFTTAGYSVNGTTGAVT 402

Qy        284 AGI-------AQAIATAGLVQAYLPGKSMMAIGGG 311
```

```
                |:        :|: |:  |  ||     :  | |
Db        403 KGVDSVYVDNNEALTTSDTVDFYLQDDGSVTNGSG 437


RESULT 5
PMP8_CHLPN
ID   PMP8_CHLPN     STANDARD;      PRT;   930 AA.
AC   Q9Z393; Q9RB66;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Probable outer membrane protein pmp8 precursor (Polymorphic membrane
DE   protein 8) (Outer membrane protein 11).
GN   PMP8 OR OMP11 OR CPN0446 OR CP0307.
OS   Chlamydia pneumoniae (Chlamydophila pneumoniae).
OC   Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX   NCBI_TaxID=83558;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=VR1310;
RX   MEDLINE=20007584; PubMed=10539856;
RA   Christiansen G., Boesen T., Hjerno K., Daugaard L., Mygind P.,
RA   Madsen A.S., Knudsen K., Falk E., Birkelund S.;
RT   "Molecular biology of Chlamydia pneumoniae surface proteins and their
RT   role in immunopathogenicity.";
RL   Am. Heart J. 138:S491-S495(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CWL029;
RX   MEDLINE=99206606; PubMed=10192388;
RA   Kalman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,
RA   Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
RT   "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";
RL   Nat. Genet. 21:385-389(1999).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AR39;
RX   MEDLINE=20150255; PubMed=10684935;
RA   Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA   White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,
RA   Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA   Gwinn M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
RA   Eisen J., Fraser C.M.;
RT   "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
RT   pneumoniae AR39.";
RL   Nucleic Acids Res. 28:1397-1406(2000).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=J138;
RX   MEDLINE=20330349; PubMed=10871362;
RA   Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
RA   Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
RT   "Comparison of whole genome sequences of Chlamydia pneumoniae J138
RT   from Japan and CWL029 from USA.";
RL   Nucleic Acids Res. 28:2311-2314(2000).
CC   -!- SUBCELLULAR LOCATION: CELL WALL SURFACE (ELEMENTARY BODIES)
CC       (POTENTIAL).
CC   -!- SIMILARITY: BELONGS TO THE PMP OUTER MEMBRANE PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AJ133034; CAB37068.1; -.
DR   EMBL; AE001627; AAD18590.1; -.
DR   EMBL; AE002193; AAF38164.1; -.
DR   EMBL; AP002546; BAA98654.1; -.
DR   PIR; A81591; A81591.
DR   PIR; D72078; D72078.
```

```
DR   PHCI-2DPAGE; Q9Z393; -.
DR   TIGR; CP0307; -.
DR   InterPro; IPR006315; Autotransport.
DR   InterPro; IPR005546; Autotransporter.
DR   InterPro; IPR003368; Chlamydia_PMP.
DR   Pfam; PF03797; Autotransporter; 1.
DR   Pfam; PF02415; DUF145; 2.
DR   TIGRFAMs; TIGR01414; autotrans_barl; 1.
DR   TIGRFAMs; TIGR01376; POMP_repeat; 6.
KW   Outer membrane; Signal; Multigene family; Complete proteome.
FT   SIGNAL        1     26       POTENTIAL.
FT   CHAIN        27    930       PROBABLE OUTER MEMBRANE PROTEIN PMP8.
FT   CONFLICT    177    177       T -> A (IN REF. 3 AND 4).
SQ   SEQUENCE   930 AA;  97669 MW;  46A9B5E3BB913C4C CRC64;

  Query Match             7.5%;  Score 136.5;  DB 1;  Length 930;
  Best Local Similarity   22.6%;  Pred. No. 1.1;
  Matches   88;  Conservative   48;  Mismatches  156;  Indels   97;  Gaps   18;

Qy         11 TVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKDKGE-NGSSTDEG 69
              |  |  |   | |    :  ||  |   :     :     | :||    |  |  |  |
Db         34 TDSFDGAGGSTFTPKSTADANG--TNYVLSGNVYINDAGKGTALTGCCFTETTGDLTFTG 91

Qy         70 EGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKGTTATVSKD- 128
              :|   :   :|| : ||    ||         |||  | |  :|::| :  |||    |
Db         92 KGYSFSFNTVDAGSNAGAAASTT--------ADKALTFTGFSNLSFIAAPGTTVASGKST 143

Qy        129 ----------DQGNITVMYDVNV-------GDALNVNQLQNSGWNLDSKAVAGSSGKVIS 171
                     | |  |::::  ||       | |:    |  || |  |     :| | :
Db        144 LSSAGALNLTDNG--TILFSQNVSNEANNNGGAITTKTLSISG-NTSSITFTSNSAKKLG 200

Qy        172 GNVSPSKGKMDETVNINAGNNIEITRNGK------NIDIATSMTPQFSSV---------S 216
              | :  |      ::: | |  : :   |:        : ::|:| | ||:         :
Db        201 GAIYSSAA---ASISGNTGQLVFMNNKGETGGGALGFEASSSIT-QNSSLFFSGNTATDA 256

Qy        217 LGAG--------ADAPTLSVDGDALNVGSKKDNKPVRITNVAPGVKEGDV-TNVAQLKGV 267
               | |         : |||:: |          || :     :   : | :  :   |
Db        257 AGKGGAIYCEKTGETPTLTISG----------NKSLTFAENSSVTQGGAICAHGLDLSAA 306

Qy        268 AQNL--NNRIDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSS 325
                  |  |||     ||  ||   ||| |                    |     :
Db        307 GPTLFSNNRC----GNTAAGKGGAIAIAD------------------SGSLSLSANQGD 343

Qy        326 ISDGGNWIIKGTASGNSRG--HFGASASV 352
              |:  || :   :|  ::|   : |:|| :
Db        344 ITFLGNTLTSTSAPTSTRNAIYLGSSAKI 372


RESULT 6
120K_RICRI
ID   120K_RICRI     STANDARD;      PRT;  1300 AA.
AC   P14914;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   120 kDa surface-exposed protein.
GN   P120.
OS   Rickettsia rickettsii.
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC   Rickettsiaceae; Rickettsieae; Rickettsia.
OX   NCBI_TaxID=783;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=R;
RX   MEDLINE=90136087; PubMed=2515418;
RA   Gilmore R.D. Jr., Joste N., McDonald G.A.;
RT   "Cloning, expression and sequence analysis of the gene encoding the
RT   120 kD surface-exposed protein of Rickettsia rickettsii.";
RL   Mol. Microbiol. 3:1579-1586(1989).
CC   -!- FUNCTION: MAJOR STRUCTURAL PROTEIN WHICH MAY PLAY A ROLE AS
CC       RICKETTSIAL VIRULENCE FACTOR AND/OR IMMUNOGEN DURING INFECTION.
CC   -!- SUBCELLULAR LOCATION: CELL WALL. THIS BACTERIUM IS COVERED BY A
CC       S-LAYER WITH HEXAGONAL SYMMETRY.
CC   -!- MISCELLANEOUS: ITS C-TERMINUS POTENTIALLY MAY BEAR THE EPITOPES
CC       CONFERRING ANTIGENICITY TO THE PROTEIN.
CC   -!- SIMILARITY: BELONGS TO THE RICKETTSIAE OMPA/OMPB FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X16353; CAA34402.1; -.
DR   PIR; S07575; S07575.
DR   InterPro; IPR006315; Autotransport.
DR   InterPro; IPR005546; Autotransporter.
DR   Pfam; PF03797; Autotransporter; 1.
DR   TIGRFAMs; TIGR01414; autotrans_barl; 2.
KW   Antigen; Glycoprotein; Cell wall; S-layer.
FT   CARBOHYD      7      7       N-LINKED (POTENTIAL).
FT   CARBOHYD     66     66       N-LINKED (POTENTIAL).
FT   CARBOHYD     86     86       N-LINKED (POTENTIAL).
FT   CARBOHYD    103    103       N-LINKED (POTENTIAL).
FT   CARBOHYD    147    147       N-LINKED (POTENTIAL).
FT   CARBOHYD    268    268       N-LINKED (POTENTIAL).
FT   CARBOHYD    330    330       N-LINKED (POTENTIAL).
FT   CARBOHYD    375    375       N-LINKED (POTENTIAL).
FT   CARBOHYD    415    415       N-LINKED (POTENTIAL).
FT   CARBOHYD    424    424       N-LINKED (POTENTIAL).
FT   CARBOHYD    430    430       N-LINKED (POTENTIAL).
FT   CARBOHYD    436    436       N-LINKED (POTENTIAL).
FT   CARBOHYD    444    444       N-LINKED (POTENTIAL).
FT   CARBOHYD    515    515       N-LINKED (POTENTIAL).
FT   CARBOHYD    547    547       N-LINKED (POTENTIAL).
FT   CARBOHYD    593    593       N-LINKED (POTENTIAL).
FT   CARBOHYD    655    655       N-LINKED (POTENTIAL).
FT   CARBOHYD    698    698       N-LINKED (POTENTIAL).
FT   CARBOHYD    710    710       N-LINKED (POTENTIAL).
FT   CARBOHYD    799    799       N-LINKED (POTENTIAL).
FT   CARBOHYD    800    800       N-LINKED (POTENTIAL).
FT   CARBOHYD    826    826       N-LINKED (POTENTIAL).
FT   CARBOHYD    844    844       N-LINKED (POTENTIAL).
FT   CARBOHYD    861    861       N-LINKED (POTENTIAL).
FT   CARBOHYD    879    879       N-LINKED (POTENTIAL).
FT   CARBOHYD    920    920       N-LINKED (POTENTIAL).
FT   CARBOHYD    926    926       N-LINKED (POTENTIAL).
FT   CARBOHYD   1116   1116       N-LINKED (POTENTIAL).
FT   CARBOHYD   1128   1128       N-LINKED (POTENTIAL).
FT   CARBOHYD   1140   1140       N-LINKED (POTENTIAL).
FT   CARBOHYD   1146   1146       N-LINKED (POTENTIAL).
FT   CARBOHYD   1211   1211       N-LINKED (POTENTIAL).
SQ   SEQUENCE   1300 AA;  132801 MW;  E09E52C3F647243D CRC64;

  Query Match             7.5%;  Score 136.5;  DB 1;  Length 1300;
  Best Local Similarity   22.6%;  Pred. No. 1.6;
  Matches   95;  Conservative   47;  Mismatches  162;  Indels  117;  Gaps   19;

Qy          2 NVDFVRTYD---TVEFLSADTKTTTVN----VESKDNG-------KKTEVKIGAKTSVIK 47
              |   | |:|   |:|  |||       |    :|:   |           |:::|   |: |
Db         76 NTAGVITFDANGTLESASADANVAVTNNITAIEASGAGVVQLSGTHAAELRLGNAGSIFK 135

Qy         48 EKDGKLVTGKDK------GENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQA 101
                || :: ||         |   : |   :|  |     |      |     : | ||
Db        136 LADGTVINGKVNQTALVGGALAAGTITLDGSATITGDIGNAGGAAALQRITLAN------ 189

Qy        102 DKFETVT-SGTNVTFASG-------KGTTATVSKDDQGNITVMYDV-------NVGDALN 146
              |  :|:|  | |:  | |         | |  ::   | || | :|:        | || :
Db        190 DAKKTLTLGGANIIGAGGGTIDLQANGGTIKLT-STQNNIVVDFDLAIATDQTGVVDASS 248

Qy        147 VNQLQNSGWN-------LDSKAVA----GSSGKVIS-GNV-------------------- 174
              :   |    |         ::| :     |||  |:| |||
Db        249 LTNAQTLTINGKIGTIGANNKTLGQFNIGSSKTVLSNGNVAINELVIGNDGAVQFAHDTY 308
```

```
Qy        175 -------SPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLS 227
                     :  :||:     :| |  :     | |:  ||:   : :  | |   |
Db        309 LITRTTNAAGQGKIIFNPVVNNGTTLAA---GTNLGSATNPLAEINFGSKGVNVDT---- 361

Qy        228 VDGDALNVGSKKDNKPVRIT----NVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNAR 283
                   ||||   :     ||    ||   |     ||:     |     |: : |
Db        362 ----VLNVGEGVNLYATNITTTDANVGSFVFNAGGTNIVS-GTVGGQQGNKFNTV----- 411

Qy        284 AGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSR 343
                        |   | ::  :|  |: |    |   |:: |||:     || :
Db        412 -------------ALENGTTVKFLGNATFNGNTTIAAN-STLQIGGNYTADCVASADGT 456

Qy        344 G 344
               |
Db        457 G 457


RESULT 7
OMPB_RICRI
ID   OMPB_RICRI     STANDARD;      PRT;  1654 AA.
AC   Q53047;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Outer membrane protein B precursor (168 kDa surface-layer protein)
DE   (Surface protein antigen) (Cell surface antigen 5) (Sca5) (rOmpB)
DE   (rOmp B) [Contains: 120 kDa surface-exposed protein (Surface protein
DE   antigen) (120 kDa outer membrane protein ompB); 32 kDa beta peptide].
GN   OMPB.
OS   Rickettsia rickettsii.
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC   Rickettsiaceae; Rickettsieae; Rickettsia.
OX   NCBI_TaxID=783;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=R;
RX   MEDLINE=92167802; PubMed=1724278;
RA   Gilmore R.D. Jr., Cieplak W. Jr., Policastro P.F., Hackstadt T.;
RT   "The 120 kilodalton outer membrane protein (rOmp B) of Rickettsia
RT   rickettsii is encoded by an unusually long open reading frame:
RT   evidence for protein processing from a large precursor.";
RL   Mol. Microbiol. 5:2361-2370(1991).
RN   [2]
RP   SEQUENCE OF 279-1654 FROM N.A.
RC   STRAIN=R;
RX   MEDLINE=90136087; PubMed=2515418;
RA   Gilmore R.D. Jr., Joste N., McDonald G.A.;
RT   "Cloning, expression and sequence analysis of the gene encoding the
RT   120 kD surface-exposed protein of Rickettsia rickettsii.";
RL   Mol. Microbiol. 3:1579-1586(1989).
CC   -!- FUNCTION: THE 120 kDa SURFACE-EXPOSED PROTEIN IS A MAJOR
CC       STRUCTURAL PROTEIN WHICH MAY PLAY A ROLE AS A RICKETTSIAL
CC       VIRULENCE FACTOR AND/OR IMMUNOGEN DURING INFECTION.
CC   -!- SUBCELLULAR LOCATION: CELL WALL. THIS BACTERIUM IS COVERED BY A S-
CC       LAYER WITH HEXAGONAL SYMMETRY.
CC   -!- SIMILARITY: BELONGS TO THE RICKETTSIAE OMPA/OMPB FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X16353; CAA34403.1; -.
DR   PIR; S18227; S18227.
DR   InterPro; IPR006315; Autotransport.
DR   InterPro; IPR005546; Autotransporter.
DR   Pfam; PF03797; Autotransporter; 1.
DR   TIGRFAMs; TIGR01414; autotrans_barl; 2.
KW   Antigen; S-layer; Cell wall.
FT   CHAIN        1   1333       120 kDa SURFACE-EXPOSED PROTEIN.
FT   CHAIN     1334   1654       32 kDa BETA PEPTIDE.
FT   DOMAIN    1181   1188       POLY-THR.
SQ   SEQUENCE   1654 AA;  168184 MW;  D7AB70FB7087F618 CRC64;

  Query Match             7.5%;  Score 136.5;  DB 1;  Length 1654;
  Best Local Similarity   22.6%;  Pred. No. 2.1;
  Matches   95;  Conservative   47;  Mismatches  162;  Indels  117;  Gaps   19;

Qy          2 NVDFVRTYD---TVEFLSADTKTTTVN----VESKDNG-------KKTEVKIGAKTSVIK 47
              |   | |:|   |:|  |||       |    :|:   |         |:::|   |: |
Db        430 NTAGVITFDANGTLESASADANVAVTNNITAIEASGAGVVQLSGTHAAELRLGNAGSIFK 489

Qy         48 EKDGKLVTGKDK------GENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQA 101
                || :: ||        |   : |   :|  |    |     |    : | ||
Db        490 LADGTVINGKVNQTALVGGALAAGTITLDGSATITGDIGNAGGAAALQRITLAN------ 543

Qy        102 DKFETVT-SGTNVTFASG-------KGTTATVSKDDQGNITVMYDV-------NVGDALN 146
              |  :|:|  | |:  | |        | |  ::   | || | :|:        | || :
Db        544 DAKKTLTLGGANIIGAGGGTIDLQANGGTIKLT-STQNNIVVDFDLAIATDQTGVVDASS 602

Qy        147 VNQLQNSGWN-------LDSKAVA----GSSGKVIS-GNV-------------------- 174
              :   |    |        ::| :     |||  |:| |||
Db        603 LTNAQTLTINGKIGTIGANNKTLGQFNIGSSKTVLSNGNVAINELVIGNDGAVQFAHDTY 662

Qy        175 -------SPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLS 227
                    :  :||:     :| |  :     | |:  ||:   : :  | |   |
Db        663 LITRTTNAAGQGKIIFNPVVNNGTTLAA---GTNLGSATNPLAEINFGSKGVNVDT---- 715

Qy        228 VDGDALNVGSKKDNKPVRIT----NVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNAR 283
                   ||||   :     ||    ||   |     ||:     |     |: : |
Db        716 ----VLNVGEGVNLYATNITTTDANVGSFVFNAGGTNIVS-GTVGGQQGNKFNTV----- 765

Qy        284 AGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSR 343
                        |   | ::  :|  |: |    |   |:: |||:     || :
Db        766 -------------ALENGTTVKFLGNATFNGNTTIAAN-STLQIGGNYTADCVASADGT 810

Qy        344 G 344
               |
Db        811 G 811


RESULT 8
FLIC_ECOLI
ID   FLIC_ECOLI     STANDARD;      PRT;   497 AA.
AC   P04949;
DT   13-AUG-1987 (Rel. 05, Created)
DT   13-AUG-1987 (Rel. 05, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Flagellin.
GN   FLIC OR FLAF OR HAG OR B1923.
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12;
RX   MEDLINE=87057066; PubMed=3536885;
RA   Kuwajima G., Asaka J.-I., Fujiwara T., Fujiwara T., Node K., Kondo E.;
RT   "Nucleotide sequence of the hag gene encoding flagellin of
RT   Escherichia coli.";
RL   J. Bacteriol. 168:1479-1483(1986).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12;
RX   MEDLINE=89281489; PubMed=2659972;
RA   Hanafusa T., Sakai A., Tominaga A., Enomoto M.;
RT   "Isolation and characterization of Escherichia coli hag operator
RT   mutants whose hag48 expression has become repressible by a Salmonella
RT   H1 repressor.";
RL   Mol. Gen. Genet. 216:44-50(1989).
RN   [3]
RP   SEQUENCE FROM N.A.
```

```
RC   STRAIN=K12 / MG1655;
RX   MEDLINE=97426617; PubMed=9278503;
RA   Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA   Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA   Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA   Mau B., Shao Y.;
RT   "The complete genome sequence of Escherichia coli K-12.";
RL   Science 277:1453-1474(1997).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12;
RX   MEDLINE=97251358; PubMed=9097040;
RA   Itoh T., Aiba H., Baba T., Fujita K., Hayashi K., Inada T.,
RA   Isono K., Kasai H., Kimura S., Kitakawa M., Kitagawa M.,
RA   Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,
RA   Nakade S., Nakamura Y., Nashimoto H., Nishio Y., Oshima T.,
RA   Saito N., Sampei G., Seki Y., Sivasundaram S., Tagami H.,
RA   Takeda J., Takemoto K., Wada C., Yamamoto Y., Horiuchi T.;
RT   "A 460-kb DNA sequence of the Escherichia coli K-12 genome
RT   corresponding to the 40.1-50.0 min region on the linkage map.";
RL   DNA Res. 3:379-392(1996).
RN   [5]
RP   SEQUENCE OF 1-19 FROM N.A.
RX   MEDLINE=83238225; PubMed=6305924;
RA   Szekely E., Simon M.;
RT   "DNA sequence adjacent to flagellar genes and evolution of flagellar-
RT   phase variation.";
RL   J. Bacteriol. 155:74-81(1983).
RN   [6]
RP   SEQUENCE OF 1-12.
RC   STRAIN=K12 / EMG2;
RX   MEDLINE=97443975; PubMed=9298646;
RA   Link A.J., Robison K., Church G.M.;
RT   "Comparing the predicted and observed properties of proteins encoded
RT   in the genome of Escherichia coli K-12.";
RL   Electrophoresis 18:1259-1313(1997).
RN   [7]
RP   SEQUENCE OF 1-4.
RC   STRAIN=K12 / W3110;
RX   MEDLINE=98263247; PubMed=9600841;
RA   Wilkins M.R., Gasteiger E., Tonella L., Ou K., Tyler M.,
RA   Sanchez J.-C., Gooley A.A., Walsh B.J., Bairoch A., Appel R.D.,
RA   Williams K.L., Hochstrasser D.F.;
RT   "Protein identification with N and C-terminal sequence tags in
RT   proteome projects.";
RL   J. Mol. Biol. 278:599-608(1998).
CC   -!- FUNCTION: FLAGELLIN IS THE SUBUNIT PROTEIN WHICH POLYMERIZES TO
CC       FORM THE FILAMENTS OF BACTERIAL FLAGELLA.
CC   -!- SIMILARITY: BELONGS TO THE BACTERIAL FLAGELLIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M14358; AAA23950.1; -.
DR   EMBL; X17440; CAA35488.1; -.
DR   EMBL; AE000285; AAC74990.1; -.
DR   EMBL; D90832; BAA15744.1; -.
DR   EMBL; D90833; BAA15751.1; -.
DR   EMBL; J01607; AAA92491.1; -.
DR   PIR; A37249; FLEC.
DR   SWISS-2DPAGE; P04949; COLI.
DR   EcoGene; EG10321; fliC.
DR   InterPro; IPR001029; Flagellin_C.
DR   InterPro; IPR001492; FlagellinN.
DR   Pfam; PF00700; Flagellin_C; 1.
DR   Pfam; PF00669; Flagellin_N; 1.
DR   PRINTS; PR00207; FLAGELLIN.
KW   Flagella; Complete proteome.
FT   INIT_MET      0      0
FT   CONFLICT    284    284       P -> L (IN REF. 2).
SQ   SEQUENCE   497 AA;  51163 MW;  24B5419C21C7B4E8 CRC64;

  Query Match             7.5%;  Score 135.5;  DB 1;  Length 497;
  Best Local Similarity   21.1%;  Pred. No. 0.59;
  Matches   78;  Conservative   66;  Mismatches  147;  Indels   79;  Gaps   15;

Qy         10 DTVEPLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEG 69
              | :: :|  |:   ||| :|:   |  :::||  :     | | :  |  | :| |
Db        120 DEIDRVSGQTQFNGVNVLAKNGSMK--IQVGANDNQTITIDLKQIDAKTLGLDGFSVKNN 177

Qy         70 EGLVTAKEV-------IDAVNKAGWRMKT--TTANGQTGQA-----------DKFETVTS 109
              : : |:  |         : :   |  : |   |  | |  |                | :  :|
Db        178 DTVTTSAPVTAFGATTTNNIKLTGITLSTEAATDTGGTNPASIEGVYTDNGNDYYAKITG 237

Qy        110 GTN--------------VTFASGKGTTATVSKDD--------QGNITVMYDVNVGD---- 143
              | |              || |:|    |||:  :          |   |  |  |
Db        238 GDNDGKYYAVTVANDGTVTMATGATANATVTDANTTKATTITSGGTPVQIDNTAGSATAN 297

Qy        144 --ALNVNQLQNS-GWNLDSKAVAGSSGKVISGNVSPSKGKM--------DETVNINAGNN 192
                |::: :||:| | : |: |:  ::| : : :|: : | :         | :   ::
Db        298 LGAVSLVKLQDSKGNDTDTYALKDTNGNLYAADVNETTGAVSVKTITYTDSSGAASSPTA 357

Qy        193 IEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGDALN-VGSKKDNKPVRITNVAP 251
              :::  :   ::       : |  |  |      |:  |:||  | : |   |::  : |
Db        358 VKLGGDDGKTEVVDIDGKTYDSADLNGGNLQTGLTAGGEALTAVANGKTTDPLKALDDAI 417

Qy        252 GVKE------GDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQA--------IATAGLVQ 297
                :      | | |  :|     ||||   |:   |:: |  |         :: | ::|
Db        418 ASVDKFRSSLGAVQN--RLDSAVTNLNNTTTNL-SEAQSRIQDADYATEVSNMSKAQIIQ 474

Qy        298 AYLPGKSMMA 307
                | |::|
Db        475 Q--AGNSVLA 482


RESULT 9
AIDA_ECOLI
ID   AIDA_ECOLI     STANDARD;      PRT;  1286 AA.
AC   Q03155;
DT   01-JUN-1994 (Rel. 29, Created).
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Adhesin aidA-I precursor.
GN   AIDA-I.
OS   Escherichia coli.
OG   Plasmid pIB6.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 50-56.
RC   STRAIN=O126:H27 / 2787;
RX   MEDLINE=92326638; PubMed=1625582;
RA   Benz I., Schmidt M.A.;
RT   "AIDA-I, the adhesin involved in diffuse adherence of the
RT   diarrhoeagenic Escherichia coli strain 2787 (O126:H27), is
RT   synthesized via a precursor molecule.";
RL   Mol. Microbiol. 6:1539-1546(1992).
CC   -!- FUNCTION: THIS IS AN ADHESION PROTEIN NECESSARY FOR THE DIFFUSE
CC       ADHERENCE OF DIARRHEA-CAUSING ENTEROPATHOGENIC ESCHERICHIA COLI
CC       TO EPITHELIAL CELLS.
CC   -!- SUBCELLULAR LOCATION: Outer membrane.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
```

```
DR   EMBL; X65022; CAA46156.1; -.
DR   PIR; S28634; S28634.
DR   InterPro; IPR006315; Autotransport.
DR   InterPro; IPR005546; Autotransporter.
DR   InterPro; IPR004899; Pertactin.
DR   Pfam; PF03797; Autotransporter; 1.
DR   Pfam; PF03212; Pertactin; 1.
DR   TIGRFAMs; TIGR01414; autotrans_barl; 2.
KW   Cell adhesion; Signal; Outer membrane; Plasmid.
FT   SIGNAL       1     49
FT   CHAIN       50      ?       ADHESIN AIDA-I.
FT   PROPEP       ?   1286
SQ   SEQUENCE   1286 AA;  132271 MW;  B2A00F72AC05FB34 CRC64;

  Query Match             7.4%;  Score 134.5;  DB 1;  Length 1286;
  Best Local Similarity   24.4%;  Pred. No. 2;
  Matches   89;  Conservative   48;  Mismatches  145;  Indels   83;  Gaps   19;

Qy         64 SSTDEGEGLVTAKEVI-----------DAVNKAGW-RMKTTTANGQT-----GQADKFE 105
              |    | | | ||  :             ||| :|      | ::|:|     |: :
Db         20 SELARGHGFVLAKNTLLVLAVVSTIGNAFAVNISGTVSSGGTVSSGETQIVYSGRGNSNA 79

Qy        106 TVTS-GTNVTFASGKGTTATVSKDDQGNI-----TVMYDVNVGDALNVN--------QLQ 151
              || | || :    || |  ||:     |:     |:   || |    |:         |
Db         80 TVNSGGTQIVNNGGKTTATTVNSSGSQNVGTSGATISTIVNSGGIQRVSSGGVASATNLS 139

Qy        152 NSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQ 211
                  |: :   | ::     || :   | : :: ||::|    ::  |    :|:: |
Db        140 GGAQNIYNLGHASNTVIFSGGNQTIFSGGITDSTNISSGGQQRVSSGG----VASNTTIN 195

Qy        212 FS---SVSLGAGADAPTLSVDGDA-LNVGSKKDNKPVRITNVA--PGVKEGDVTNVAQLK 265
               |   ::    || :  :|  |:  :: |:   |    | |       |  | |     |
Db        196 SSGAQNILSEEGAISTHISSGGNQYISAGA---NATETIVNSGGFQRVNSGAVATGTVLS 252

Qy        266 GVAQNL-----------NNRIDNV-------DGNARAGIAQAIATAGLV---------- 296
               |  ||:            |: :  |       |    :|  | |:: |:|
Db        253 GGTQNVSSGGSAISTSVYNSGVQTVFAGATVTDTTVNSGGNQNISSGGIVSETTVNVSGT 312

Qy        297 -QAYLPGKSMMA-IGGGTYRGEAGYAIGYSSISDGGNWIIK--GTASG---NSRGHFGAS 349
                 |  | :: | | |       | ||  : :||||   |:  | |||   |  |:   |
Db        313 QNIYSGGSALSANIKGSQIVNSEGTAIN-TLVSDGGYQHIRNGGIASGTIVNQSGYVNIS 371

Qy        350 ASVGY 354
               | ||
Db        372 -SGGY 375


RESULT 10
HLYA_SERMA
ID   HLYA_SERMA     STANDARD;      PRT;  1608 AA.
AC   P15320;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   01-NOV-1990 (Rel. 16, Last annotation update)
DE   Hemolysin precursor.
GN   SHLA.
OS   Serratia marcescens.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Serratia.
OX   NCBI_TaxID=615;
RN   [1]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 31-40.
RC   STRAIN=SN8;
RX   MEDLINE=88257037; PubMed=3290200;
RA   Poole K., Schiebel E., Braun V.;
RT   "Molecular characterization of the hemolysin determinant of Serratia
RT   marcescens.";
RL   J. Bacteriol. 170:3177-3188(1988).
CC   -!- FUNCTION: BACTERIAL HEMOLYSINS ARE EXOTOXINS THAT ATTACK BLOOD
CC       CELL MEMBRANES AND CAUSE CELL RUPTURE BY MECHANISMS NOT CLEARLY
CC       DEFINED.
CC   -!- FUNCTION: CELL-BOUND HEMOLYSIN, WHICH RELEASES HEME-IRON FROM
CC       ERYTHROCYTES BY INTERACTION WITH THE ERYTHROCYTE MEMBRANE. SHLA
CC       REQUIRES SHLB FUNCTION.
CC   -!- SUBCELLULAR LOCATION: Outer membrane.
CC   -!- SIMILARITY: TO P.MIRABILIS HEMOLYSIN (HPMA).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M22618; AAA50323.1; -.
DR   PIR; A28182; A28182.
KW   Hemolysis; Toxin; Outer membrane; Signal.
FT   SIGNAL       1     30
FT   CHAIN       31   1608       HEMOLYSIN.
SQ   SEQUENCE   1608 AA;  165078 MW;  D669B476FE7DAD51 CRC64;

  Query Match             7.3%;  Score 132.5;  DB 1;  Length 1608;
  Best Local Similarity   22.7%;  Pred. No. 3.3;
  Matches   98;  Conservative   49;  Mismatches  159;  Indels  125;  Gaps   20;

Qy         15 LSADTKTTTVNVE-----------------------------------SKDNGKKTEV 37
              || ::||  :||:                                   ::|: | |
Db        668 LSVESKTGNINVKAAERQQNIDEQKTALTVNGYAKEAGDKQYRAGLRIEHTRDSEKTTRT 727

Qy         38 KIGAKTSVIKEKDGKLVTGKDKGENGSS--TDEGEGLVTAKEV--IDAVNKAGWRMKTTT 93
              :  |  | :     ||   ||   :||    |:|:  |:  :|  : | :|     : |
Db        728 ENSA--SSLSGGSVKLKAEKDVTFSGSKLVADKGDASVSGNKVSFLAADDKTASNTEQTK 785

Qy         94 ANG---QTGQADKFETVTSGTNVTFASGKGTTATVSK------DDQGNITVMYDVNVGDA 144
                 |    ||  ||   : ||    : : | | |  ||      | :||:|    :|  |
Db        786 IGGGFYYTGGIDK---LGSGVEAGYENNK-TQAQSSKAITSGSDVKGNLT----INARDK 837

Qy        145 LNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDI 204
              |     |:|      :  ||      :   | :  | |  |||           | |:|
Db        838 LTQQGAQHSVGGAYQENAAGVDHLAAADTASTTTTKTDVGVNI----------GANVDY 886

Qy        205 ATSMTPQFSSVSLGA-----------GADAPTLSVDGDALNVGSKK--DNKPVRITNVA 250
              :   |   :|   |            |  || : :|  |    |:|   :    :::|
Db        887 SAVTRPVERAVGKAAKLDATGVINDIGGIGAPNVGLDIGAQGGSSEKRSSSSQAVVSSVQ 946

Qy        251 PGV----KEGDV----TNVAQLKGVAQNLN----------NRIDNVDGNARAGIAQAIAT 292
              |      :|:|    |     || | ||           || |    : |       :
Db        947 AGSIDINAKGEVRDQGTQYQASKG-AVNLTADSHRSEAAANRQDEQSRDTR-------GS 998

Qy        293 AGLVQAYLPGKSMMAIG----GGTYRGEAGYAIGYSSISDGGNW---------IIKGTAS 339
              || |: |    | : :     ||| |  :  :   :   |  |           | :|||
Db        999 AG-VRVYTTTGSDLTVDAKGEGGTQRSNSSASQAVTGSIDAANGINVNVKKDAIYQGTAL 1057

Qy        340 GNSRGHFGASA 350
                 ||    :|
Db       1058 NGGRGKTAVNA 1068


RESULT 11
YADA_YERPS
ID   YADA_YERPS     STANDARD;      PRT;   434 AA.
AC   P10858;
DT   01-JUL-1989 (Rel. 11, Created)
DT   01-JUL-1989 (Rel. 11, Last sequence update)
DT   01-JUL-1993 (Rel. 26, Last annotation update)
DE   Invasin precursor (Outer membrane adhesin).
GN   YADA OR YOPA OR INVA OR YOP1.
OS   Yersinia pseudotuberculosis.
OG   Plasmid pIB1.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Yersinia.
OX   NCBI_TaxID=633;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=YPIII;
```

```
RX   MEDLINE=88302441; PubMed=3043229;
RA   Rosqvist R., Skurnik M., Wolf-Watz H.;
RT   "Increased virulence of Yersinia pseudotuberculosis by two
RT   independent mutations.";
RL   Nature 334:522-525(1988).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=YPIII;
RX   MEDLINE=89343638; PubMed=2761389;
RA   Skurnik M., Wolf-Watz H.;
RT   "Analysis of the yopA gene encoding the Yop1 virulence determinants
RT   of Yersinia spp.";
RL   Mol. Microbiol. 3:517-529(1989).
CC   -!- FUNCTION: INVASIN IS A PROTEIN THAT ALLOWS ENTERIC BACTERIA TO
CC       PENETRATE CULTURED MAMMALIAN CELLS. THE ENTRY OF INVASIN IN THE
CC       CELL IS MEDIATED BY BINDING SEVERAL BETA-1 CHAIN INTEGRINS. THIS
CC       PROTEIN IS AN ADHESIN FORMING A FIBRILLAR MATRIX ON THE CELL
CC       SURFACE.
CC   -!- SUBCELLULAR LOCATION: Outer membrane.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X13883; CAA32088.1; -.
DR   PIR; S04534; S04534.
DR   InterPro; IPR005594; YadA.
DR   Pfam; PF03895; YadA; 1.
KW   Plasmid; Virulence; Signal; Outer membrane.
FT   SIGNAL       1     25
FT   CHAIN       26    434       INVASIN.
SQ   SEQUENCE   434 AA;  45054 MW;  EE2C55FB12B183D4 CRC64;

  Query Match             7.2%;  Score 131;  DB 1;  Length 434;
  Best Local Similarity   21.1%;  Pred. No. 0.91;
  Matches   84;  Conservative   51;  Mismatches  137;  Indels  126;  Gaps   16;

Qy         24 VNVESKDNGKKTEVKIG-----AKTSVIKEKDGKLVTG----------KDKGEN-----G 63
              :|  :||      : ||     || : :    | : ||          |  |::
Db         98 LNARAKD---PYSIAIGATAEAAKPAAVAVGSGSIATGVNSVAIGPLSKALGDSAVTYGA 154

Qy         64 SSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKGTTA 123
              ||| : :|:       |:       |  | |   : |   :|  | :   |:  | :
Db        155 SSTAQKDGV--------AIGARASASDTGVAVGFNSKVDAQNSVAIGHSSHVAADHGYSI 206

Qy        124 TV---SKDDQGNITVMYDVNVG-DALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKG 179
               :   || |: |      |::| ::||      :    |: ||         ||:  |
Db        207 AIGDHSKTDREN-----SVSIGHESLNRQLTHLAAGTEDTDAV----------NVAQLKK 251

Qy        180 KMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGDALNVGSKK 239
              :| ||:                          |   ||:   | |:   |
Db        252 EMAETLE--------------------------------NARKETLAQSNDVLDAAKKH 278

Qy        240 DNKPVRIT-------------------------NVAPGVKEG----DVT-NVAQLKGVAQ 269
               |   | |                        | :  :|      ||| : :  | :::
Db        279 SNSVARTTLETAEEHANKKSAEALVSAKVYADSNSSHTLKTANSYTDVTVSSSTKKAISE 338

Qy        270 N----------LNNRIDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGY 319
              :          |:||:| :|     |:| : |   | | |  ||     | | ||
Db        339 SNQYTDHKFSQLDNRLDKLDKRVDKGLASSAALNSLFQPYGVGKVNFTAGVGGYRSSQAL 398

Qy        320 AIGYSSISDGGNWIIKGTA-SGNSRGHFGASASVGYQW 356
              |||     :   : | | :|:|   : || ::  :|
Db        399 AIGSGYRVNESVALKAGVAYAGSSNVMYNASFNI--EW 434


RESULT 12
TOP2_PLAFK
ID   TOP2_PLAFK     STANDARD;      PRT;  1398 AA.
AC   P41001;
DT   01-FEB-1995 (Rel. 31, Created)
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   DNA topoisomerase II (EC 5.99.1.3).
GN   TOP2.
OS   Plasmodium falciparum (isolate K1 / Thailand).
OC   Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX   NCBI_TaxID=5839;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94316496; PubMed=8041616;
RA   Cheesman S., McAleese S., Goman M., Johnson D., Horrocks P.,
RA   Ridley R.G., Kilbey B.J.;
RT   "The gene encoding topoisomerase II from Plasmodium falciparum.";
RL   Nucleic Acids Res. 22:2547-2551(1994).
CC   -!- FUNCTION: CONTROL OF TOPOLOGICAL STATES OF DNA BY TRANSIENT
CC       BREAKAGE AND SUBSEQUENT REJOINING OF DNA STRANDS. TOPOISOMERASE II
CC       MAKES DOUBLE-STRAND BREAKS.
CC   -!- CATALYTIC ACTIVITY: ATP-dependent breakage, passage and rejoining
CC       of double-stranded DNA.
CC   -!- SUBUNIT: Homodimer (By similarity).
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- MISCELLANEOUS: EUKARYOTIC TOPOISOMERASE I AND II CAN RELAX BOTH
CC       NEGATIVE AND POSITIVE SUPERCOILS, WHEREAS PROKARYOTIC ENZYMES
CC       RELAX ONLY NEGATIVE SUPERCOILS.
CC   -!- SIMILARITY: Belongs to the type II topoisomerase family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X79345; -; NOT_ANNOTATED_CDS.
DR   HSSP; P06786; 1BGW.
DR   InterPro; IPR003594; ATPbind_ATPase.
DR   InterPro; IPR003957; CBFA_NFYB_topis.
DR   InterPro; IPR001241; DNA_topoisoII.
DR   InterPro; IPR002205; DNA_topoisoIV.
DR   Pfam; PF00204; DNA_gyraseB; 1.
DR   Pfam; PF00521; DNA_topoisoIV; 1.
DR   Pfam; PF02518; HATPase_c; 1.
DR   PRINTS; PR00615; CCAATSUBUNTA.
DR   PRINTS; PR00418; TPI2FAMILY.
DR   ProDom; PD000742; DNA_topoisoIV; 1.
DR   SMART; SM00387; HATPase_c; 1.
DR   SMART; SM00433; TOP2c; 1.
DR   SMART; SM00434; TOP4c; 1.
DR   PROSITE; PS00177; TOPOISOMERASE_II; 1.
KW   Isomerase; Topoisomerase; DNA-binding; ATP-binding; Nuclear protein.
FT   NP_BIND     144    149       ATP (POTENTIAL).
FT   ACT_SITE    830    830       DNA CLEAVAGE (BY SIMILARITY).
FT   DOMAIN      271    281       POLY-ASN.
FT   DOMAIN      308    316       POLY-ASN.
FT   DOMAIN     1089   1093       POLY-LYS.
FT   DOMAIN     1227   1234       POLY-LYS.
SQ   SEQUENCE   1398 AA;  161029 MW;  BAAD7BEE88FE5BE9 CRC64;

  Query Match             7.2%;  Score 131;  DB 1;  Length 1398;
  Best Local Similarity   24.2%;  Pred. No. 3.5;
  Matches   54;  Conservative   46;  Mismatches   89;  Indels   34;  Gaps    9;

Qy          4 DFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAK---TSVIKEKDGKLVTGKDKG 60
              |  :  : :||           |||  :  :  : |:  |   :|: |:|  | ::  ::
Db       1190 DIEKVEEAIEFQR--------NVELSNREESNKFKVARKQGPSSMKKKKKKKKLSSDEES 1241

Qy         61 ENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKG 120
              | | ::|  | ||     :    ||     ||||::     : |        |   ::
Db       1242 EGGDTSDSSEFLVNTLNIKKNTNK-----KTTTSSNNVNNSKKRLRKADDLN---SNELD 1293
```

```
Qy        121 TTATVSK--DDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGK---VISGNVS 175
              | :|||  ||  |:|   :  : : ||    : |  |:|:|:   :| |    |: : :
Db       1294 NTLSVSKTFDDNNNLT--DNTPLINRLNDENNEFSSNNVDNKSTNKNSRKKKPKIADSTN 1351

Qy        176 PSKGKMDETVNINAGNNIEITRNGKNIDIATSMT---PQFSSV 215
                  :  ::: :: ||   | :|     || |: : |    :|||:
Db       1352 DNNSELNSSIQINDNVNDDI-----NITISPNKTINVNEFSSI 1389


RESULT 13
YPJA_ECOLI
ID   YPJA_ECOLI     STANDARD;      PRT;  1569 AA.
AC   P52143; P76610; P77017; P77019;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Hypothetical outer membrane protein ypjA.
GN   YPJA OR B2647.
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12 / MG1655;
RX   MEDLINE=97426617; PubMed=9278503;
RA   Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA   Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA   Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA   Mau B., Shao Y.;
RT   "The complete genome sequence of Escherichia coli K-12.";
RL   Science 277:1453-1474(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12;
RX   MEDLINE=97349980; PubMed=9205837;
RA   Yamamoto Y., Aiba H., Baba T., Hayashi K., Inada T., Isono K.,
RA   Itoh T., Kimura S., Kitagawa M., Makino K., Miki T., Mitsuhashi N.,
RA   Mizobuchi K., Mori H., Nakade S., Nakamura Y., Nashimoto H.,
RA   Oshima T., Oyama S., Saito N., Sampei G., Satoh Y., Sivasundaram S.,
RA   Tagami H., Takahashi H., Takeda J., Takemoto K., Uehara K., Wada C.,
RA   Yamagata S., Horiuchi T.;
RT   "Construction of a contiguous 874-kb sequence of the Escherichia coli
RT   - K12 genome corresponding to 50.0-68.8 min on the linkage map and
RT   analysis of its sequence features.";
RL   DNA Res. 4:91-113(1997).
CC   -!- SUBCELLULAR LOCATION: Outer membrane (Potential).
CC   -!- SIMILARITY: STRONG, TO BORDETELLA PERTACTIN.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U36840; AAA79815.1; ALT_SEQ.
DR   EMBL; AE000350; AAC75695.1; -.
DR   EMBL; D90889; BAA16514.1; ALT_INIT.
DR   EMBL; D90890; BAA16518.1; ALT_INIT.
DR   PIR; A65044; A65044.
DR   EcoGene; EG13213; ypjA.
DR   InterPro; IPR006315; Autotransport.
DR   InterPro; IPR005546; Autotransporter.
DR   InterPro; IPR004899; Pertactin.
DR   InterPro; IPR003991; Pertactin_C.
DR   Pfam; PF03797; Autotransporter; 1.
DR   Pfam; PF03212; Pertactin; 1.
DR   PRINTS; PR01484; PRTACTNFAMLY.
DR   TIGRFAMs; TIGR01414; autotrans_barl; 2.
KW   Hypothetical protein; Outer membrane; Complete proteome.
SQ   SEQUENCE   1569 AA;  162774 MW;  0D006A9A32154596 CRC64;
```

```
  Query Match             7.2%;  Score 130.5;  DB 1;  Length 1569;
  Best Local Similarity   19.3%;  Pred. No. 4.2;
  Matches   69;  Conservative   60;  Mismatches  191;  Indels   37;  Gaps   10;

Qy          1 NNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKDKG 60
              | |   | :          | ||:     | | |  |: |||   :  :| ::  :| |
Db        523 NYVVLENTGELTVVAKTSAKNTTI-----DTGGKLIVQKEAKTDSTRLNNGGVLEVQDGG 577

Qy         61 ENGSSTDEGEGLV----TAKEVIDAVNKAG--WRMKTTTANGQTGQADKFETVTSGTNV- 113
              |      :  | :    |:  :|:  |  |  : :: : |     :     || :|:
Db        578 EAKHVEQQSGGALIASTTSGTLIEGTNSYGDAFYIRNSEAKNVVLENAGSLTVVTGSRAV 637

Qy        114 -TFASGKGTTATVSKD-----DQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSG 167
               |  :  |      ||     :      :|     |  |:   : : : | ::|   |
Db        638 DTIINANGKMDVYGKDVGTVLNSAGTQTIYASATSDKANIKGGKQTVYGLATEANIESGE 697

Qy        168 KVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLS 227
              ::: |  :       |  :|| |     :   || |: | :      :     |:   ::
Db        698 QIVDGGST-------EKTHINGGTQ-TVQNYGKAIN--TDIVSGLQQIMANGTAEGSIIN 747

Qy        228 VDGDALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIA 287
                :| |   :|  :          ::|  | : |    |     |   |:|    |:|
Db        748 GGSQVVNEGGLAENSVLNDGGTLDVREKGSATGIQQSSQGALVATTRATRVTGTRADGVA 807

Qy        288 QAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWII--KGTASGNS 342
              :|             :::   ||    |:  :   : :: ||  |:  | ||:| :
Db        808 FSIEQGA-------ANNILLANGGVLTVESDTSSDKTQVNMGGREIVKTKATATGTT 857


RESULT 14
HLYA_PROMI
ID   HLYA_PROMI     STANDARD;      PRT;  1577 AA.
AC   P16466;
DT   01-AUG-1990 (Rel. 15, Created)
DT   01-AUG-1990 (Rel. 15, Last sequence update)
DT   01-NOV-1990 (Rel. 16, Last annotation update)
DE   Hemolysin precursor.
GN   HPMA.
OS   Proteus mirabilis.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Proteus.
OX   NCBI_TaxID=584;
RN   [1]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 30-43.
RC   STRAIN=Isolate 477-12;
RX   MEDLINE=90170827; PubMed=2407716;
RA   Uphoff T.S., Welch R.A.;
RT   "Nucleotide sequencing of the Proteus mirabilis calcium-independent
RT   hemolysin genes (hpmA and hpmB) reveals sequence similarity with the
RT   Serratia marcescens hemolysin genes (shlA and shlB).";
RL   J. Bacteriol. 172:1206-1216(1990).
CC   -!- FUNCTION: BACTERIAL HEMOLYSINS ARE EXOTOXINS THAT ATTACK BLOOD
CC       CELL MEMBRANES AND CAUSE CELL RUPTURE BY MECHANISMS NOT CLEARLY
CC       DEFINED.
CC   -!- FUNCTION: CELL-BOUND HEMOLYSIN, WHICH RELEASES HEME-IRON FROM
CC       ERYTHROCYTES BY INTERACTION WITH THE ERYTHROCYTE MEMBRANE. HPMA
CC       REQUIRES HPMB FUNCTION.
CC   -!- SUBCELLULAR LOCATION: Outer membrane.
CC   -!- MISCELLANEOUS: THE CONSERVED AMPHIPATHIC DOMAINS IN SHLA AND HPMA
CC       MAY BE RESPONSIBLE FOR PORE FORMATION.
CC   -!- SIMILARITY: TO S.MARCESCENS HEMOLYSIN (SHLA).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M30186; AAA25657.1; -.
DR   PIR; A35140; A35140.
```

```
KW   Hemolysis; Toxin; Outer membrane; Signal.
FT   SIGNAL        1     29
FT   CHAIN        30   1577       HEMOLYSIN.
SQ   SEQUENCE   1577 AA;  165869 MW;  175975E0C924B2D9 CRC64;

  Query Match             7.2%;  Score 130;  DB 1;  Length 1577;
  Best Local Similarity   22.8%;  Pred. No. 4.5;
  Matches   92;  Conservative   56;  Mismatches  130;  Indels  126;  Gaps   24;

Qy          1 NNVDF---VRTYDTVEFLS--ADTKTTTVNVE----------SKDNGKKTEVKIGAKTSV 45
              || |    |   | :  |:  :: :|  | ||          |: : :     | |
Db        895 NNTDLTKKVTARDAIANLANLSNLETPNVGVEVGIKGGGSQQSQTDSQAVSTSINAGKID 954

Qy         46 IKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDA-VNKAGWRMKTTTANGQTGQADKF 104
              | : : ||    |:| :  || ||  |       :| ::|       |   || | : |
Db        955 I-DSNNKL---HDQGTHYQSTQEGISLTANTHTSEATLDKHQTTFHETKGGGQIGVSTK- 1009

Qy        105 ETVTSGTNVTFA-SGKGTTA------TVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNL 157
                  :|:::| |  |:| |       | :|  |        :|||:          |
Db       1010 ----TGSDITVAIKGEGQTTDNALMETKAKGSQFTSNGDISINVGE-----NAHYEGAQF 1060

Qy        158 DSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITR-----------NGK-NIDI 204
              |::              |||   || |||| :: : :            ||  |: :
Db       1061 DAQ----------------KGK---TV-INAGGDLTLAQATDTHSESQSNVNGSANLKV 1099

Qy        205 ATSMTPQFSSVSLGAGADAPTL--SVDGDALNVGSKKDNKPVRITNVAPGVKEGDVTNVA 262
                |  ||:  |   | | :| |   | :     ||:   :: : :        :|  |:::
Db       1100 GT--TPE--SKDYGGGFNAGTTHHSKEQTTAKVGTITGSQGIELNAGHNLTLQG--THLS 1153

Qy        263 QLKGVAQNLNNRID----------NVDGNARAGIAQ-----AIATAGLVQAYLPGKS 304
              : :| |  |::|            |: |  :||  :     | :  ||       | :
Db       1154 SEQDIALNATNKVDLQSASSEHTEKGNNLSGGVQAGFGKKMTDDASSVNGL------GSA 1207

Qy        305 MMAIG----------GGTYRGEAGYAIGYSSISDGGNWIIKGTA 338
                |||          |||             |:: ||  | | :
Db       1208 QFAIGKQDEKSVSREGGT------------INNSGNLTINGNS 1238


RESULT 15
OMPA_RICRI
ID   OMPA_RICRI     STANDARD;      PRT;  2249 AA.
AC   P15921;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Outer membrane protein A precursor (190 kDa antigen) (Cell surface
DE   antigen) (rOmpA) (rOmp A).
GN   OMPA.
OS   Rickettsia rickettsii.
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC   Rickettsiaceae; Rickettsieae; Rickettsia.
OX   NCBI_TaxID=783;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=R;
RX   MEDLINE=90354033; PubMed=2117568;
RA   Anderson B.E., McDonald G.A., Jones D.C., Regnery R.L.;
RT   "A protective protein antigen of Rickettsia rickettsii has tandemly
RT   repeated, near-identical sequences.";
RL   Infect. Immun. 58:2760-2769(1990).
CC   -!- FUNCTION: ELICITS PROTECTIVE IMMUNITY.
CC   -!- SUBCELLULAR LOCATION: CELL WALL. THIS BACTERIUM IS COVERED BY A
CC       S-LAYER WITH HEXAGONAL SYMMETRY.
CC   -!- PTM: GLYCOSYLATED (PROBABLE).
CC   -!- SIMILARITY: BELONGS TO THE RICKETTSIAE OMPA/OMPB FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M31227; AAA26380.1; -.
DR   PIR; A41477; A41477.
DR   InterPro; IPR006315; Autotransport.
DR   InterPro; IPR005546; Autotransporter.
DR   Pfam; PF03797; Autotransporter; 1.
DR   TIGRFAMs; TIGR01414; autotrans_barl; 3.
KW   Antigen; Repeat; Signal; Cell wall; S-layer; Glycoprotein.
FT   SIGNAL        1     28       POTENTIAL.
FT   CHAIN        29   2249       OUTER MEMBRANE PROTEIN A.
FT   DOMAIN      212   1180       13 X APPROXIMATE TANDEM REPEATS.
FT   REPEAT      212    286       A (TYPE I).
FT   REPEAT      287    358       B (TYPE II).
FT   REPEAT      359    430       C (TYPE II).
FT   REPEAT      431    505       D (TYPE I).
FT   REPEAT      506    577       E (TYPE II).
FT   REPEAT      578    652       F (TYPE I).
FT   REPEAT      653    724       G (TYPE II).
FT   REPEAT      725    799       H (TYPE I).
FT   REPEAT      800    874       I (TYPE I).
FT   REPEAT      875    949       J (TYPE I).
FT   REPEAT      950   1021       K (TYPE II).
FT   REPEAT     1022   1093       L (TYPE II).
FT   REPEAT     1094   1165       M (TYPE II).
FT   REPEAT     1166   1180       TYPE I (INCOMPLETE).
SQ   SEQUENCE   2249 AA;  224333 MW;  A9D6646C089DF087 CRC64;

  Query Match             7.2%;  Score 130;  DB 1;  Length 2249;
  Best Local Similarity   25.8%;  Pred. No. 6.8;
  Matches  104;  Conservative   38;  Mismatches  163;  Indels   98;  Gaps   23;

Qy         11 TVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKL--VTGKDKGENGSST-D 67
                | :  :| :  |  |  :   |        |    |:   :| |  |||     |  :|
Db        902 TTKLTNAASVLTLTNANAVLTGAIDNTTGGDNVGVL-NLNGALSQVTGDIGNTNSLATIS 960

Qy         68 EGEGLVT-----AKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKGTT 122
                | |  |      |     :  |   :| |      ||  |       :|  |||      |
Db        961 VGAGTATLGGAVIKATTTKLTDAASAVKFTNPVVVTGAIDNTGNANNGI-VTFTGNSTVT 1019

Qy        123 ATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMD 182
                |     ||   :  ||||  |   |:|      ::  :  ::  |   |     | :|
Db       1020 GNV-----GNTNALATVNVGAGL--LQVQGGVVKANTINLTDNASAVTFTNPVVVTGAID 1072

Qy        183 ETVNINAG-----NNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGDAL--NV 235
                | | | |      |  :| |  |        |   ::|::|||     | | |  :  |
Db       1073 NTGNANNGIVTFTGNSTVTGNVGN-------TNALATVNVGAG----LLQVQGGVVKANT 1121

Qy        236 GSKKDN-KPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARA----GIAQAI 290
                :  ||   |  ||  | |  | : |        | || |    ||:      |   |:
Db       1122 INLTDNASAVTFTN--PVVVTGAIDNTG-------NANNGIVTFTGNSTVTGDIGNTNAL 1172

Qy        291 AT----AGL-VQAYLPGKSMMA----------------IGGGTYRGEAGYAIGY---SSI 326
                ||    ||: :||   | |: |               : ||      | || |    :|
Db       1173 ATVNVGAGITLQA---GGSLAANNIDFGARSTLEFNGPLDGG------GKAIPYYFKGAI 1223

Qy        327 SDGGNWIIK-------------GTASGNSRGH---FGASASVG 353
                ::| | |:              || :  : |    |   ||||
Db       1224 ANGNNAILNVNTKLLTASHLTIGTVAEINIGAGNLFTIDASVG 1266


Search completed: October  6, 2003, 09:24:18
Job time : 6.49755 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: October 6, 2003, 09:13:10 ; Search time 26.1469 Seconds
(without alignments)
3513.485 Million cell updates/sec

Title: US-09-771-382-37
Perfect score: 1813
Sequence: 1 NNVDFVRTYDTVEFLSADTK.....TASGNSRGHFGASASVGYQW 356

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_23:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1808	99.7	591	2	Q93QY3	Q93qy3 neisseria m
2	1808	99.7	591	2	Q9JPS7	Q9jps7 neisseria m
3	1808	99.7	591	16	Q9JR18	Q9jr18 neisseria m
4	1804	99.5	592	2	Q9AQF0	Q9aqf0 neisseria m
5	1787	98.6	526	2	Q9JPS4	Q9jps4 neisseria m
6	1787	98.6	530	2	Q9JPS1	Q9jps1 neisseria m
7	1770	97.6	600	2	Q9JPS6	Q9jps6 neisseria m
8	1762.5	97.2	592	2	Q9JPS9	Q9jps9 neisseria m
9	1758.5	97.0	598	2	Q9JPR7	Q9jpr7 neisseria m
10	1758.5	97.0	599	2	Q9JPS8	Q9jps8 neisseria m
11	1756.5	96.9	590	2	Q9JPS3	Q9jps3 neisseria m
12	1753.5	96.7	594	2	Q93QY4	Q93qy4 neisseria m
13	1752.5	96.7	592	16	Q9JQW4	Q9jqw4 neisseria m
14	1750.5	96.6	594	2	Q9JPI3	Q9jpi3 neisseria m
15	1750.5	96.6	594	2	Q9JPS2	Q9jps2 neisseria m
16	1750.5	96.6	595	2	Q9JPH0	Q9jph0 neisseria m
17	1748.5	96.4	599	2	Q9JPR8	Q9jpr8 neisseria m
18	1744.5	96.2	600	2	Q9JPS5	Q9jps5 neisseria m
19	1741.5	96.1	594	2	Q9JPH7	Q9jph7 neisseria m
20	1741.5	96.1	598	2	Q9JPR9	Q9jpr9 neisseria m
21	1728.5	95.3	592	2	Q93QY2	Q93qy2 neisseria m
22	1728.5	95.3	598	2	Q9JPS0	Q9jps0 neisseria m
23	1728.5	95.3	598	2	Q93QY5	Q93qy5 neisseria m
24	1728.5	95.3	598	2	Q9JPT0	Q9jpt0 neisseria m
25	1727.5	95.3	589	2	Q9JPI0	Q9jpi0 neisseria m
26	1715.5	94.6	589	2	Q93QY1	Q93qy1 neisseria m
27	608	33.5	1098	2	Q48152	Q48152 haemophilus
28	606.5	33.5	1096	2	Q8GM79	Q8gm79 haemophilus
29	586.5	32.3	1210	2	Q8GM74	Q8gm74 haemophilus
30	577.5	31.9	1210	2	Q8GM75	Q8gm75 haemophilus
31	572.5	31.6	1204	2	Q8GM76	Q8gm76 haemophilus
32	569	31.4	2353	2	P71401	P71401 haemophilus
33	369	20.4	1299	16	Q9F3X6	Q9f3x6 pasteurella
34	335.5	18.5	1004	2	Q8GM77	Q8gm77 haemophilus
35	334.5	18.5	1002	2	Q8GM78	Q8gm78 haemophilus
36	332.5	18.3	2059	16	Q9PD50	Q9pd50 xylella fas
37	327	18.0	1190	16	Q9PC04	Q9pc04 xylella fas
38	325	17.9	1588	16	Q8XDG4	Q8xdg4 escherichia
39	318	17.5	1778	16	Q8FCB2	Q8fcb2 escherichia
40	315	17.4	1461	16	Q8ZL64	Q8zl64 salmonella
41	314	17.3	1107	16	Q9F2D8	Q9f2d8 salmonella
42	310.5	17.1	2314	2	Q8KQM8	Q8kqm8 moraxella c
43	307.5	17.0	641	16	Q8CKM1	Q8ckm1 yersinia pe
44	307.5	17.0	658	16	Q8ZHJ0	Q8zhj0 yersinia pe
45	295.5	16.3	2712	16	Q9F3X5	Q9f3x5 pasteurella

ALIGNMENTS

RESULT 1
Q93QY3
ID Q93QY3 PRELIMINARY; PRT; 591 AA.
AC Q93QY3;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE NhhA outer membrane protein.
GN NHHA.
OS Neisseria meningitidis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=487;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EG329;
RA Peak I.R., Srikhanta Y., Dieckelman M., Moxon R., Jennings M.P.;
RT "Identification and characterization of a gene encoding a novel outer
RT membrane protein of Neisseria meningitidis.";
RL Submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF157606; AAK68867.1; -.
DR InterPro; IPR005594; YadA.
DR Pfam; PF03895; YadA; 1.
SQ SEQUENCE 591 AA; 62048 MW; C0DC600798859C65 CRC64;

Query Match 99.7%; Score 1808; DB 2; Length 591;
Best Local Similarity 99.7%; Pred. No. 2.3e-81;
Matches 355; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 NNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKDKG 60
              :||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        236 DNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKDKG 295

Qy         61 ENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        296 ENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKG 355

Qy        121 TTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGK 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        356 TTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGK 415

Qy        181 MDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGDALNVGSKKD 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        416 MDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGDALNVGSKKD 475

Qy        241 NKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAYL 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        476 NKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAYL 535

Qy        301 PGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 356
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        536 PGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 591


RESULT 2
Q9JPS7
ID   Q9JPS7      PRELIMINARY;      PRT;   591 AA.
AC   Q9JPS7;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Outer membrane protein GNA992.
GN   GNA992.
OS   Neisseria meningitidis.
OC   Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC   Neisseriaceae; Neisseria.
OX   NCBI_TaxID=487;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BZ147;
RX   MEDLINE=20175756; PubMed=10710308;
RA   Pizza M., Scarlato V., Masignani V., Giuliani M.M., Arico' B.,
RA   Comanducci M., Jennings G.T., Baldi L., Bartolini E., Capecchi B.,
RA   Galeotti C.L., Luzzi E., Manetti R., Marchetti E., Mora M., Nuti S.,
RA   Ratti G., Santini L., Savino S., Scarselli M., Storni E., Zuo P.,
RA   Broeker M., Hundt E., Knapp B., Blair E., Mason T., Tettelin H.,
RA   Hood D.W., Jeffries A.C., Saunders N.J., Granoff D.M., Venter C.,
RA   Moxon E.R., Grandi G., Rappuoli R.;
RT   "Identification of Vaccine Candidates Against Serogroup B
RT   Meningococcus by Whole-Genome Sequencing.";
RL   Science 287:1816-1820(2000).
DR   EMBL; AF226366; AAF42515.1; -.
DR   InterPro; IPR005594; YadA.
DR   Pfam; PF03895; YadA; 1.
SQ   SEQUENCE   591 AA;  62113 MW;  533453CAE5A91E1F CRC64;

  Query Match             99.7%;  Score 1808;  DB 2;  Length 591;
  Best Local Similarity   99.7%;  Pred. No. 2.3e-81;
  Matches  355;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 NNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKDKG 60
              :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        236 DNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKDKG 295

Qy         61 ENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        296 ENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKG 355

Qy        121 TTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGK 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        356 TTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGK 415

Qy        181 MDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGDALNVGSKKD 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        416 MDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGDALNVGSKKD 475

Qy        241 NKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAYL 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        476 NKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAYL 535

Qy        301 PGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 356
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        536 PGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 591


RESULT 3
Q9JR18
ID   Q9JR18      PRELIMINARY;      PRT;   591 AA.
AC   Q9JR18;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Outer membrane protein GNA992 (Adhesin) (NhhA outer membrane
DE   protein).
GN   GNA992 OR NMB0992 OR NHHA.
OS   Neisseria meningitidis, and
OS   Neisseria meningitidis (serogroup B).
OC   Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC   Neisseriaceae; Neisseria.
OX   NCBI_TaxID=487, 491;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MC58 / Serogroup B, BZ169, BZ83, and H44/76;
RX   MEDLINE=20175756; PubMed=10710308;
RA   Pizza M., Scarlato V., Masignani V., Giuliani M.M., Arico' B.,
RA   Comanducci M., Jennings G.T., Baldi L., Bartolini E., Capecchi B.,
RA   Galeotti C.L., Luzzi E., Manetti R., Marchetti E., Mora M., Nuti S.,
RA   Ratti G., Santini L., Savino S., Scarselli M., Storni E., Zuo P.,
RA   Broeker M., Hundt E., Knapp B., Blair E., Mason T., Tettelin H.,
RA   Hood D.W., Jeffries A.C., Saunders N.J., Granoff D.M., Venter C.,
RA   Moxon E.R., Grandi G., Rappuoli R.;
RT   "Identification of Vaccine Candidates Against Serogroup B
RT   Meningococcus by Whole-Genome Sequencing.";
RL   Science 287:1816-1820(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MC58 / Serogroup B;
RX   MEDLINE=20175755; PubMed=10710307;
RA   Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
RA   Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
RA   Nelson W.C., Gwinn M.L., DeBoy R., Peterson J.D., Hickey E.K.,
RA   Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
RA   Mason T., Ciecko A., Parksey D.S., Blair E., Cittone H., Clark E.B.,
RA   Cotton M.D., Utterback T.R., Khouri H., Qin H., Vamathevan J.,
RA   Gill J., Scarlato V., Masignani V., Pizza M., Grandi G., Sun L.,
RA   Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;
RT   "Complete genome sequence of Neisseria meningitidis serogroup B strain
RT   MC58.";
RL   Science 287:1809-1815(2000).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   SPECIES=N.meningitidis; STRAIN=PMC21;
RA   Peak I.R., Srikhanta Y., Dieckelman M., Moxon R., Jennings M.P.;
RT   "Identification and characterization of a gene encoding a novel outer
RT   membrane protein of Neisseria meningitidis.";
RL   Submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF226375; AAF42524.1; -.
DR   EMBL; AE002450; AAF41395.1; -.
DR   EMBL; AF226367; AAF42516.1; -.
DR   EMBL; AF226370; AAF42519.1; -.
DR   EMBL; AF226374; AAF42523.1; -.
DR   EMBL; AF157611; AAK68872.1; -.
DR   TIGR; NMB0992; -.
DR   InterPro; IPR005594; YadA.
DR   Pfam; PF03895; YadA; 1.
KW   Complete proteome.
SQ   SEQUENCE   591 AA;  62112 MW;  7C22F3CAE7F73EC6 CRC64;

  Query Match             99.7%;  Score 1808;  DB 16;  Length 591;
  Best Local Similarity   99.7%;  Pred. No. 2.3e-81;
  Matches  355;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 NNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKDKG 60
              :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        236 DNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKDKG 295
```

```
Qy          61 ENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKG 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         296 ENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKG 355

Qy         121 TTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGK 180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         356 TTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGK 415

Qy         181 MDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGDALNVGSKKD 240
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         416 MDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGDALNVGSKKD 475

Qy         241 NKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAYL 300
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         476 NKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAYL 535

Qy         301 PGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 356
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         536 PGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 591
```

RESULT 4
Q9AQF0
ID Q9AQF0 PRELIMINARY; PRT; 592 AA.
AC Q9AQF0;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Outer membrane protein.
GN NHHA.
OS Neisseria meningitidis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=487;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MC58;
RA Peak I.R., Srikhanta Y., Dieckelman M., Moxon E.R., Jennings M.P.;
RT "Identification and characterisation of a gene encoding a novel outer
RT membrane protein of Neisseria meningitidis.";
RL Submitted (FEB-1999) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF125375; AAK09243.1; -.
DR InterPro; IPR005594; YadA.
DR Pfam; PF03895; YadA; 1.
SQ SEQUENCE 592 AA; 62290 MW; 168986A97381EFC5 CRC64;

Query Match 99.5%; Score 1804; DB 2; Length 592;
Best Local Similarity 99.4%; Pred. No. 3.7e-81;
Matches 354; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
Qy           1 NNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKDKG 60
               :|||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db         237 DNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGVKTSVIKEKDGKLVTGKDKG 296

Qy          61 ENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKG 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         297 ENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKG 356

Qy         121 TTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGK 180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         357 TTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGK 416

Qy         181 MDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGDALNVGSKKD 240
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         417 MDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGDALNVGSKKD 476

Qy         241 NKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAYL 300
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         477 NKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAYL 536

Qy         301 PGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 356
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         537 PGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 592
```

RESULT 5
Q9JPS4
ID Q9JPS4 PRELIMINARY; PRT; 526 AA.
AC Q9JPS4;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Outer membrane protein GNA992.
GN GNA992.
OS Neisseria meningitidis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=487;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NG6/88;
RX MEDLINE=20175756; PubMed=10710308;
RA Pizza M., Scarlato V., Masignani V., Giuliani M.M., Arico' B.,
RA Comanducci M., Jennings G.T., Baldi L., Bartolini E., Capecchi B.,
RA Galeotti C.L., Luzzi E., Manetti R., Marchetti E., Mora M., Nuti S.,
RA Ratti G., Santini L., Savino S., Scarselli M., Storni E., Zuo P.,
RA Broeker M., Hundt E., Knapp B., Blair E., Mason T., Tettelin H.,
RA Hood D.W., Jeffries A.C., Saunders N.J., Granoff D.M., Venter C.,
RA Moxon E.R., Grandi G., Rappuoli R.;
RT "Identification of Vaccine Candidates Against Serogroup B
RT Meningococcus by Whole-Genome Sequencing.";
RL Science 287:1816-1820(2000).
DR EMBL; AF226377; AAF42526.1; -.
DR InterPro; IPR005594; YadA.
DR Pfam; PF03895; YadA; 1.
SQ SEQUENCE 526 AA; 54732 MW; F543C4234AF807CA CRC64;

Query Match 98.6%; Score 1787; DB 2; Length 526;
Best Local Similarity 98.6%; Pred. No. 2.2e-80;
Matches 351; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

```
Qy           1 NNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKDKG 60
               :||||| ||||||||||||||||||||||||:|||||||||||||||||||||| ||
Db         171 DNVDFVHTYDTVEFLSADTKTTTVNVESKDNGKRTEVKIGAKTSVIKEKDGKLVTGKGKG 230

Qy          61 ENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKG 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         231 ENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKG 290

Qy         121 TTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGK 180
               ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db         291 TTATVSKDDQGNITVKYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGK 350

Qy         181 MDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGDALNVGSKKD 240
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         351 MDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGDALNVGSKKD 410

Qy         241 NKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAYL 300
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         411 NKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAYL 470

Qy         301 PGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 356
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         471 PGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 526
```

RESULT 6
Q9JPS1
ID Q9JPS1 PRELIMINARY; PRT; 530 AA.
AC Q9JPS1;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Outer membrane protein GNA992.
GN GNA992.

```
OS   Neisseria meningitidis.
OC   Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC   Neisseriaceae; Neisseria.
OX   NCBI_TaxID=487;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NGF26;
RX   MEDLINE=20175756; PubMed=10710308;
RA   Pizza M., Scarlato V., Masignani V., Giuliani M.M., Arico' B.,
RA   Comanducci M., Jennings G.T., Baldi L., Bartolini E., Capecchi B.,
RA   Galeotti C.L., Luzzi E., Manetti R., Marchetti E., Mora M., Nuti S.,
RA   Ratti G., Santini L., Savino S., Scarselli M., Storni E., Zuo P.,
RA   Broeker M., Hundt E., Knapp B., Blair E., Mason T., Tettelin H.,
RA   Hood D.W., Jeffries A.C., Saunders N.J., Granoff D.M., Venter C.,
RA   Moxon E.R., Grandi G., Rappuoli R.;
RT   "Identification of Vaccine Candidates Against Serogroup B
RT   Meningococcus by Whole-Genome Sequencing.";
RL   Science 287:1816-1820(2000).
DR   EMBL; AF226380; AAF42529.1; -.
DR   InterPro; IPR005594; YadA.
DR   Pfam; PF03895; YadA; 1.
SQ   SEQUENCE   530 AA;  55190 MW;  1F836CA57598515B CRC64;

  Query Match             98.6%;  Score 1787;  DB 2;  Length 530;
  Best Local Similarity   98.6%;  Pred. No. 2.2e-80;
  Matches  351;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 NNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKDKG 60
              :|||||| |||||||||||||||||||||||:|||||||||||||||||||||||| ||
Db        175 DNVDFVHTYDTVEFLSADTKTTTVNVESKDNGKRTEVKIGAKTSVIKEKDGKLVTGKGKG 234

Qy         61 ENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKG 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        235 ENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKG 294

Qy        121 TTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGK 180
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        295 TTATVSKDDQGNITVKYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGK 354

Qy        181 MDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGDALNVGSKKD 240
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        355 MDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGDALNVGSKKD 414

Qy        241 NKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAYL 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        415 NKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAYL 474

Qy        301 PGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 356
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        475 PGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 530


RESULT 7
Q9JPS6
ID   Q9JPS6      PRELIMINARY;      PRT;   600 AA.
AC   Q9JPS6;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Outer membrane protein GNA992.
GN   GNA992.
OS   Neisseria meningitidis.
OC   Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC   Neisseriaceae; Neisseria.
OX   NCBI_TaxID=487;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=E26;
RX   MEDLINE=20175756; PubMed=10710308;
RA   Pizza M., Scarlato V., Masignani V., Giuliani M.M., Arico' B.,
RA   Comanducci M., Jennings G.T., Baldi L., Bartolini E., Capecchi B.,
RA   Galeotti C.L., Luzzi E., Manetti R., Marchetti E., Mora M., Nuti S.,
RA   Ratti G., Santini L., Savino S., Scarselli M., Storni E., Zuo P.,
RA   Broeker M., Hundt E., Knapp B., Blair E., Mason T., Tettelin H.,
RA   Hood D.W., Jeffries A.C., Saunders N.J., Granoff D.M., Venter C.,
RA   Moxon E.R., Grandi G., Rappuoli R.;
RT   "Identification of Vaccine Candidates Against Serogroup B
RT   Meningococcus by Whole-Genome Sequencing.";
RL   Science 287:1816-1820(2000).
DR   EMBL; AF226371; AAF42520.1; -.
DR   InterPro; IPR005594; YadA.
DR   Pfam; PF03895; YadA; 1.
SQ   SEQUENCE   600 AA;  62762 MW;  36256963E0598CD1 CRC64;

  Query Match             97.6%;  Score 1770;  DB 2;  Length 600;
  Best Local Similarity   98.0%;  Pred. No. 1.8e-79;
  Matches  349;  Conservative    1;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 NNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKDKG 60
              :|||||| |||||||||||||||||||||||||||||||||||||||||||||||| ||
Db        245 DNVDFVHTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKGKG 304

Qy         61 ENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKG 120
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        305 ENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQKGQADKFETVTSGTNVTFASGKG 364

Qy        121 TTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGK 180
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        365 TTATVSKDDQGNITVKYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGK 424

Qy        181 MDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGDALNVGSKKD 240
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        425 MDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGDALNVGSKDA 484

Qy        241 NKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAYL 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        485 NKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAYL 544

Qy        301 PGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 356
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        545 PGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 600


RESULT 8
Q9JPS9
ID   Q9JPS9      PRELIMINARY;      PRT;   592 AA.
AC   Q9JPS9;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Outer membrane protein GNA992.
GN   GNA992.
OS   Neisseria meningitidis.
OC   Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC   Neisseriaceae; Neisseria.
OX   NCBI_TaxID=487;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=860800;
RX   MEDLINE=20175756; PubMed=10710308;
RA   Pizza M., Scarlato V., Masignani V., Giuliani M.M., Arico' B.,
RA   Comanducci M., Jennings G.T., Baldi L., Bartolini E., Capecchi B.,
RA   Galeotti C.L., Luzzi E., Manetti R., Marchetti E., Mora M., Nuti S.,
RA   Ratti G., Santini L., Savino S., Scarselli M., Storni E., Zuo P.,
RA   Broeker M., Hundt E., Knapp B., Blair E., Mason T., Tettelin H.,
RA   Hood D.W., Jeffries A.C., Saunders N.J., Granoff D.M., Venter C.,
RA   Moxon E.R., Grandi G., Rappuoli R.;
RT   "Identification of Vaccine Candidates Against Serogroup B
RT   Meningococcus by Whole-Genome Sequencing.";
RL   Science 287:1816-1820(2000).
DR   EMBL; AF226361; AAF42510.1; -.
DR   InterPro; IPR005594; YadA.
DR   Pfam; PF03895; YadA; 1.
SQ   SEQUENCE   592 AA;  61917 MW;  4A3471514FD3C879 CRC64;

  Query Match             97.2%;  Score 1762.5;  DB 2;  Length 592;
```

```
  Best Local Similarity   98.0%;  Pred. No. 4e-79;
  Matches  350;  Conservative    1;  Mismatches    5;  Indels    1;  Gaps    1;

Qy          1 NNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKDKG 60
              :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db        236 DNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKGKG 295

Qy         61 ENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKG 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        296 ENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKG 355

Qy        121 TTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGK 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        356 TTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGK 415

Qy        181 MDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDG-DALNVGSKK 239
              |||||||||||||||||||||||||||||||||||||||||||||||||   ||||||| |
Db        416 MDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDKGALNVGSKD 475

Qy        240 DNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAY 299
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        476 ANKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAY 535

Qy        300 LPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 356
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        536 LPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 592


RESULT 9
Q9JPR7
ID   Q9JPR7      PRELIMINARY;      PRT;   598 AA.
AC   Q9JPR7;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Outer membrane protein GNA992.
GN   GNA992.
OS   Neisseria meningitidis.
OC   Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC   Neisseriaceae; Neisseria.
OX   NCBI_TaxID=487;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SWZ107;
RX   MEDLINE=20175756; PubMed=10710308;
RA   Pizza M., Scarlato V., Masignani V., Giuliani M.M., Arico' B.,
RA   Comanducci M., Jennings G.T., Baldi L., Bartolini E., Capecchi B.,
RA   Galeotti C.L., Luzzi E., Manetti R., Marchetti E., Mora M., Nuti S.,
RA   Ratti G., Santini L., Savino S., Scarselli M., Storni E., Zuo P.,
RA   Broeker M., Hundt E., Knapp B., Blair E., Mason T., Tettelin H.,
RA   Hood D.W., Jeffries A.C., Saunders N.J., Granoff D.M., Venter C.,
RA   Moxon E.R., Grandi G., Rappuoli R.;
RT   "Identification of Vaccine Candidates Against Serogroup B
RT   Meningococcus by Whole-Genome Sequencing.";
RL   Science 287:1816-1820(2000).
DR   EMBL; AF226385; AAF42534.1; -.
DR   InterPro; IPR005594; YadA.
DR   Pfam; PF03895; YadA; 1.
SQ   SEQUENCE   598 AA;  62431 MW;  0881CC094F33B4D4 CRC64;

  Query Match            97.0%;  Score 1758.5;  DB 2;  Length 598;
  Best Local Similarity   98.0%;  Pred. No. 6.4e-79;
  Matches  349;  Conservative    1;  Mismatches    5;  Indels    1;  Gaps    1;

Qy          2 NVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKDKGE 61
              |||||||||||||||||||||||||||||:|||||||||||||||||||||||| |||
Db        243 NVDFVRTYDTVEFLSADTKTTTVNVESKDNGKRTEVKIGAKTSVIKEKDGKLVTGKGKGE 302

Qy         62 NGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKGT 121
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 NGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKGT 362

Qy        122 TATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKM 181
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 TATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKM 422

Qy        182 DETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDG-DALNVGSKKD 240
              |||||||||||||||||||||||||||||||||||||||||||||||   ||||||| ||
Db        423 DETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDKGALNVGSKDA 482

Qy        241 NKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAYL 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        483 NKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAYL 542

Qy        301 PGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 356
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        543 PGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 598


RESULT 10
Q9JPS8
ID   Q9JPS8      PRELIMINARY;      PRT;   599 AA.
AC   Q9JPS8;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Outer membrane protein GNA992.
GN   GNA992.
OS   Neisseria meningitidis.
OC   Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC   Neisseriaceae; Neisseria.
OX   NCBI_TaxID=487;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A22;
RX   MEDLINE=20175756; PubMed=10710308;
RA   Pizza M., Scarlato V., Masignani V., Giuliani M.M., Arico' B.,
RA   Comanducci M., Jennings G.T., Baldi L., Bartolini E., Capecchi B.,
RA   Galeotti C.L., Luzzi E., Manetti R., Marchetti E., Mora M., Nuti S.,
RA   Ratti G., Santini L., Savino S., Scarselli M., Storni E., Zuo P.,
RA   Broeker M., Hundt E., Knapp B., Blair E., Mason T., Tettelin H.,
RA   Hood D.W., Jeffries A.C., Saunders N.J., Granoff D.M., Venter C.,
RA   Moxon E.R., Grandi G., Rappuoli R.;
RT   "Identification of Vaccine Candidates Against Serogroup B
RT   Meningococcus by Whole-Genome Sequencing.";
RL   Science 287:1816-1820(2000).
DR   EMBL; AF226364; AAF42513.1; -.
DR   InterPro; IPR005594; YadA.
DR   Pfam; PF03895; YadA; 1.
SQ   SEQUENCE   599 AA;  62693 MW;  46C2E974AF7F78E9 CRC64;

  Query Match            97.0%;  Score 1758.5;  DB 2;  Length 599;
  Best Local Similarity   98.0%;  Pred. No. 6.4e-79;
  Matches  349;  Conservative    1;  Mismatches    5;  Indels    1;  Gaps    1;

Qy          2 NVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKDKGE 61
              |||||||||||||||||||||||||||||:|||||||||||||||||||||||| |||
Db        244 NVDFVRTYDTVEFLSADTKTTTVNVESKDNGKRTEVKIGAKTSVIKEKDGKLVTGKGKGE 303

Qy         62 NGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKGT 121
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        304 NGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKGT 363

Qy        122 TATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKM 181
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        364 TATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKM 423

Qy        182 DETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDG-DALNVGSKKD 240
              |||||||||||||||||||||||||||||||||||||||||||||||   ||||||| ||
Db        424 DETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDKGALNVGSKDA 483

Qy        241 NKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAYL 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        484 NKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAYL 543

Qy        301 PGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 356
```

```
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        544 PGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 599


RESULT 11
Q9JPS3
ID   Q9JPS3      PRELIMINARY;      PRT;   590 AA.
AC   Q9JPS3;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Outer membrane protein GNA992.
GN   GNA992.
OS   Neisseria meningitidis.
OC   Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC   Neisseriaceae; Neisseria.
OX   NCBI_TaxID=487;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NGE28;
RX   MEDLINE=20175756; PubMed=10710308;
RA   Pizza M., Scarlato V., Masignani V., Giuliani M.M., Arico' B.,
RA   Comanducci M., Jennings G.T., Baldi L., Bartolini E., Capecchi B.,
RA   Galeotti C.L., Luzzi E., Manetti R., Marchetti E., Mora M., Nuti S.,
RA   Ratti G., Santini L., Savino S., Scarselli M., Storni E., Zuo P.,
RA   Broeker M., Hundt E., Knapp B., Blair E., Mason T., Tettelin H.,
RA   Hood D.W., Jeffries A.C., Saunders N.J., Granoff D.M., Venter C.,
RA   Moxon E.R., Grandi G., Rappuoli R.;
RT   "Identification of Vaccine Candidates Against Serogroup B
RT   Meningococcus by Whole-Genome Sequencing.";
RL   Science 287:1816-1820(2000).
DR   EMBL; AF226378; AAF42527.1; -.
DR   InterPro; IPR005594; YadA.
DR   Pfam; PF03895; YadA; 1.
SQ   SEQUENCE   590 AA;  61661 MW;  8AA476AC300D80C8 CRC64;

  Query Match             96.9%;  Score 1756.5;  DB 2;  Length 590;
  Best Local Similarity   97.5%;  Pred. No. 7.9e-79;
  Matches  348;  Conservative    3;  Mismatches    5;  Indels    1;  Gaps    1;

Qy          1 NNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKDKG 60
              :|||||||||||||||||||||||||||||||:|||||||||||||||||||||||| ||
Db        234 DNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKRTEVKIGAKTSVIKEKDGKLVTGKGKG 293

Qy         61 ENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        294 ENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKG 353

Qy        121 TTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGK 180
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        354 TTATVSKDDQGNITVKYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGK 413

Qy        181 MDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGD-ALNVGSKK 239
              |||||||||||||||||||||||||||||||||||||||||||||||| : |||||||
Db        414 MDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGALNVGSKD 473

Qy        240 DNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAY 299
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        474 ANKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAY 533

Qy        300 LPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 356
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        534 LPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 590


RESULT 12
Q93QY4
ID   Q93QY4      PRELIMINARY;      PRT;   594 AA.
AC   Q93QY4;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   NhhA outer membrane protein.
GN   NHHA.
OS   Neisseria meningitidis.
OC   Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC   Neisseriaceae; Neisseria.
OX   NCBI_TaxID=487;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=EG327;
RA   Peak I.R., Srikhanta Y., Dieckelman M., Moxon R., Jennings M.P.;
RT   "Identification and characterization of a gene encoding a novel outer
RT   membrane protein of Neisseria meningitidis.";
RL   Submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF157605; AAK68866.1; -.
DR   InterPro; IPR005594; YadA.
DR   Pfam; PF03895; YadA; 1.
SQ   SEQUENCE   594 AA;  62297 MW;  9DDD48B04B3A8EA2 CRC64;

  Query Match             96.7%;  Score 1753.5;  DB 2;  Length 594;
  Best Local Similarity   97.2%;  Pred. No. 1.1e-78;
  Matches  347;  Conservative    4;  Mismatches    5;  Indels    1;  Gaps    1;

Qy          1 NNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKDKG 60
              :|||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||
Db        238 DNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKRTEVKIGAKTSVIKEKDGKLVTGKDKG 297

Qy         61 ENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKG 120
              || ||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        298 ENDSSTDKGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKG 357

Qy        121 TTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGK 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        358 TTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGK 417

Qy        181 MDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGD-ALNVGSKK 239
              |||||||||||||||||||||||||||||||||||||||||||||||| : |||||||
Db        418 MDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGALNVGSKD 477

Qy        240 DNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAY 299
              |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db        478 ANKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNHIDNVDGNARAGIAQAIATAGLVQAY 537

Qy        300 LPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 356
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        538 LPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 594


RESULT 13
Q9JQW4
ID   Q9JQW4      PRELIMINARY;      PRT;   592 AA.
AC   Q9JQW4;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Putative surface fibril protein (Outer membrane protein GNA992).
GN   NMA1200 OR GNA992.
OS   Neisseria meningitidis (serogroup A), and
OS   Neisseria meningitidis.
OC   Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC   Neisseriaceae; Neisseria.
OX   NCBI_TaxID=65699, 487;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Z2491 / Serogroup A / Serotype 4A;
RX   MEDLINE=20222556; PubMed=10761919;
RA   Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,
RA   Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
RA   Davies R.M., Davis P., Devlin K., Feltwell T., Hamlin N., Holroyd S.,
RA   Jagels K., Leather S., Moule S., Mungall K., Quail M.A.,
RA   Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,
RA   Whitehead S., Spratt B.G., Barrell B.G.;
RT   "Complete DNA sequence of a serogroup A strain of Neisseria
RT   meningitidis Z2491.";
RL   Nature 404:502-506(2000).
```

```
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=205900, BZ133, F6124, AND Z2491;
RX   MEDLINE=20175756; PubMed=10710308;
RA   Pizza M., Scarlato V., Masignani V., Giuliani M.M., Arico' B.,
RA   Comanducci M., Jennings G.T., Baldi L., Bartolini E., Capecchi B.,
RA   Galeotti C.L., Luzzi E., Manetti R., Marchetti E., Mora M., Nuti S.,
RA   Ratti G., Santini L., Savino S., Scarselli M., Storni E., Zuo P.,
RA   Broeker M., Hundt E., Knapp B., Blair E., Mason T., Tettelin H.,
RA   Hood D.W., Jeffries A.C., Saunders N.J., Granoff D.M., Venter C.,
RA   Moxon E.R., Grandi G., Rappuoli R.;
RT   "Identification of Vaccine Candidates Against Serogroup B
RT   Meningococcus by Whole-Genome Sequencing.";
RL   Science 287:1816-1820(2000).
DR   EMBL; AL162755; CAB84461.1; -.
DR   EMBL; AF226357; AAF42506.1; -.
DR   EMBL; AF226365; AAF42514.1; -.
DR   EMBL; AF226373; AAF42522.1; -.
DR   EMBL; AF226386; AAF42535.1; -.
DR   InterPro; IPR005594; YadA.
DR   Pfam; PF03895; YadA; 1.
KW   Complete proteome.
SQ   SEQUENCE   592 AA;  61745 MW;  AD22E2F5EEF8F754 CRC64;

  Query Match             96.7%;  Score 1752.5;  DB 16;  Length 592;
  Best Local Similarity   97.5%;  Pred. No. 1.2e-78;
  Matches  347;  Conservative    3;  Mismatches    5;  Indels    1;  Gaps    1;

Qy          2 NVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKDKGE 61
              |||||||||||||||||||||||||||||:||||||||||||||||||||||| |||
Db        237 NVDFVRTYDTVEFLSADTKTTTVNVESKDNGKRTEVKIGAKTSVIKEKDGKLVTGKGKGE 296

Qy         62 NGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKGT 121
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        297 NGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKGT 356

Qy        122 TATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKM 181
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        357 TATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKM 416

Qy        182 DETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGD-ALNVGSKKD 240
              |||||||||||||:|||||||||||| |||||||||||||||||||| : |||||||
Db        417 DETVNINAGNNIEISRNGKNIDIATSMAPQFSSVSLGAGADAPTLSVDDEGALNVGSKDA 476

Qy        241 NKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAYL 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        477 NKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAYL 536

Qy        301 PGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 356
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        537 PGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 592


RESULT 14
Q9JPI3
ID   Q9JPI3      PRELIMINARY;      PRT;   594 AA.
AC   Q9JPI3;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Outer membrane protein GNA992.
GN   GNA992.
OS   Neisseria meningitidis.
OC   Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC   Neisseriaceae; Neisseria.
OX   NCBI_TaxID=487;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NG3/88, and BZ232;
RX   MEDLINE=20175756; PubMed=10710308;
RA   Pizza M., Scarlato V., Masignani V., Giuliani M.M., Arico' B.,
RA   Comanducci M., Jennings G.T., Baldi L., Bartolini E., Capecchi B.,
RA   Galeotti C.L., Luzzi E., Manetti R., Marchetti E., Mora M., Nuti S.,
RA   Ratti G., Santini L., Savino S., Scarselli M., Storni E., Zuo P.,
RA   Broeker M., Hundt E., Knapp B., Blair E., Mason T., Tettelin H.,
RA   Hood D.W., Jeffries A.C., Saunders N.J., Granoff D.M., Venter C.,
RA   Moxon E.R., Grandi G., Rappuoli R.;
RT   "Identification of Vaccine Candidates Against Serogroup B
RT   Meningococcus by Whole-Genome Sequencing.";
RL   Science 287:1816-1820(2000).
DR   EMBL; AF226376; AAF42525.1; -.
DR   EMBL; AF226369; AAF42518.1; -.
DR   InterPro; IPR005594; YadA.
DR   Pfam; PF03895; YadA; 1.
SQ   SEQUENCE   594 AA;  62086 MW;  1B25E03B90D04B46 CRC64;

  Query Match             96.6%;  Score 1750.5;  DB 2;  Length 594;
  Best Local Similarity   97.2%;  Pred. No. 1.6e-78;
  Matches  347;  Conservative    2;  Mismatches    7;  Indels    1;  Gaps    1;

Qy          1 NNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKDKG 60
              :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        238 DNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKDKG 297

Qy         61 ENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKG 120
              |||||||||||||||||||||||||||||||||||||||||||||||||| |||||| |
Db        298 ENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTKVTFASGNG 357

Qy        121 TTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGK 180
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        358 TTATVSKDDQGNITVKYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGK 417

Qy        181 MDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGD-ALNVGSKK 239
              |||||||||||||||||||||||||||||||||||||||||||||||||| : |||||||
Db        418 MDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGALNVGSKD 477

Qy        240 DNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAY 299
              |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        478 ANKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNHIDNVDGNARAGIAQAIATAGLVQAY 537

Qy        300 LPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 356
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        538 LPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 594


RESULT 15
Q9JPS2
ID   Q9JPS2      PRELIMINARY;      PRT;   594 AA.
AC   Q9JPS2;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Outer membrane protein GNA992.
GN   GNA992.
OS   Neisseria meningitidis.
OC   Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC   Neisseriaceae; Neisseria.
OX   NCBI_TaxID=487;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NGE31;
RX   MEDLINE=20175756; PubMed=10710308;
RA   Pizza M., Scarlato V., Masignani V., Giuliani M.M., Arico' B.,
RA   Comanducci M., Jennings G.T., Baldi L., Bartolini E., Capecchi B.,
RA   Galeotti C.L., Luzzi E., Manetti R., Marchetti E., Mora M., Nuti S.,
RA   Ratti G., Santini L., Savino S., Scarselli M., Storni E., Zuo P.,
RA   Broeker M., Hundt E., Knapp B., Blair E., Mason T., Tettelin H.,
RA   Hood D.W., Jeffries A.C., Saunders N.J., Granoff D.M., Venter C.,
RA   Moxon E.R., Grandi G., Rappuoli R.;
RT   "Identification of Vaccine Candidates Against Serogroup B
RT   Meningococcus by Whole-Genome Sequencing.";
RL   Science 287:1816-1820(2000).
DR   EMBL; AF226379; AAF42528.1; -.
DR   InterPro; IPR005594; YadA.
DR   Pfam; PF03895; YadA; 1.
SQ   SEQUENCE   594 AA;  62114 MW;  1E2A63A78F53D256 CRC64;
```

Query Match 96.6%; Score 1750.5; DB 2; Length 594;
Best Local Similarity 97.2%; Pred. No. 1.6e-78;
Matches 347; Conservative 2; Mismatches 7; Indels 1; Gaps 1;

```
Qy          1 NNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKDKG 60
              :|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        238 DNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKDKG 297

Qy         61 ENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKG 120
              |||||||||||||||||||||||||||||||||||||||||||||||||| |||||| |
Db        298 ENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTKVTFASGNG 357

Qy        121 TTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGK 180
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        358 TTATVSKDDQGNITVKYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGK 417

Qy        181 MDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGD-ALNVGSKK 239
              |||||||||||||||||||||||||||||||||||||||||||||||||| : |||||||
Db        418 MDETVNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDDEGALNVGSKD 477

Qy        240 DNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAY 299
               |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db        478 ANKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNHIDNVDGNARAGIAQAIATAGLVQAY 537

Qy        300 LPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 356
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        538 LPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 594
```

Search completed: October 6, 2003, 09:30:53
Job time : 28.1469 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: October 6, 2003, 09:05:45 ; Search time 32.8045 Seconds
(without alignments)
1848.329 Million cell updates/sec

Title: US-09-771-382-38
Perfect score: 1947
Sequence: 1 NRAASVKDVLNAGWNIKGVK.....TASGNSRGHFGASASVGYQW 382

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03:*
1: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1947	100.0	433	22	AAU06185	N. meningitidis PM
2	1941	99.7	512	22	AAU06182	N. meningitidis PM
3	1941	99.7	591	20	AAY27202	Amino acid sequenc
4	1941	99.7	591	20	AAY23741	A surface protein
5	1941	99.7	591	20	AAY23746	A surface protein
6	1941	99.7	591	22	AAU06171	N. meningitidis PM
7	1941	99.7	591	22	AAU06175	N. meningitidis EG
8	1938	99.5	591	21	AAY57045	BASB029 amino acid
9	1937	99.5	592	20	AAY23737	A surface protein
10	1891.5	97.1	502	22	AAU06186	N. meningitidis PM
11	1886.5	96.9	594	20	AAY23740	A surface protein
12	1886.5	96.9	594	21	AAY57044	BASB029 amino acid
13	1886.5	96.9	594	22	AAU06174	N. meningitidis EG
14	1881.5	96.6	599	20	AAY23743	A surface protein
15	1881.5	96.6	599	22	AAU06176	N. meningitidis H3
16	1874.5	96.3	594	20	AAY23739	A surface protein
17	1874.5	96.3	594	22	AAU06179	N. meningitidis BZ
18	1861.5	95.6	513	22	AAU06183	N. meningitidis H4
19	1861.5	95.6	592	20	AAY23744	A surface protein
20	1861.5	95.6	592	22	AAU06172	N. meningitidis H4
21	1861.5	95.6	598	20	AAY23738	A surface protein
22	1861.5	95.6	598	20	AAY23742	A surface protein
23	1861.5	95.6	598	22	AAU06177	N. meningitidis H1
24	1861.5	95.6	598	22	AAU06178	N. meningitidis BZ
25	1854.5	95.2	592	22	AAU06180	N. meningitidis Z2
26	1837.5	94.4	592	20	AAY27203	Amino acid sequenc
27	1817.5	93.3	589	20	AAY23745	A surface protein
28	1817.5	93.3	589	22	AAU06173	N. meningitidis P2
29	1810	93.0	407	22	AAU06184	N. meningitidis PM
30	1736.5	89.2	604	22	AAU06181	N. meningitidis su
31	710	36.5	1098	17	AAR99392	Haemophilus adhesi
32	690.5	35.5	1094	21	AAB23858	Haemophilus influe
33	660	33.9	2353	17	AAR99393	Haemophilus adhesi
34	660	33.9	2411	21	AAB23860	Haemophilus influe
35	600	30.8	116	21	AAB37832	Neisserial conserv
36	470.5	24.2	679	17	AAR99394	Haemophilus adhesi
37	470.5	24.2	679	21	AAB23855	Haemophilus influe
38	414.5	21.3	1002	21	AAB23854	Haemophilus influe
39	406	20.9	1004	21	AAB23857	Haemophilus influe
40	392	20.1	1104	21	AAB23856	Haemophilus influe
41	392	20.1	1104	21	AAB23859	Haemophilus influe
42	378	19.4	245	20	AAY27201	Amino acid sequenc
43	366	18.8	72	21	AAB37830	Neisserial conserv
44	356	18.3	2139	24	ABP71294	M. catarrhalis sur
45	354.5	18.2	2314	22	AAB69136	M. catarrhalis les

ALIGNMENTS

RESULT 1
AAU06185
ID AAU06185 standard; Protein; 433 AA.
XX
AC AAU06185;
XX
DT 24-OCT-2001 (first entry)
XX
DE N. meningitidis PMC21 NhhA deletion mutant #3.
XX
KW Surface antigen NhhA; meningococcal disease; meningitis vaccine;
KW mutant; mutein.
XX
OS Neisseria meningitidis strain PMC21.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Peptide 1..51
FT /label= Signal_peptide
FT Protein 52..433
FT /label= Mature_NhhA_deletion_mutant_#3
FT /note= "Predicted mature protein, specifically
FT claimed in claim 12"
XX
PN WO200155182-A1.
XX
PD 02-AUG-2001.
XX
PF 25-JAN-2001; 2001WO-AU00069.
XX
PR 25-JAN-2000; 2000US-0177917.
XX

```
PA   (UYQU ) UNIV QUEENSLAND.
XX
PI   Peak IRA,  Jennings MP;
XX
DR   WPI; 2001-488774/53.
DR   N-PSDB; AAS09175.
XX
PT   New NhhA surface antigen polypeptides and polynucleotides from
PT   Neisseria meningitidis, useful in producing vaccines for treating or
PT   preventing broad spectrum of Neisseria meningitidis -
XX
PS   Claim 12; Fig 8; 91pp; English.
XX
CC   The present invention relates to the isolation of novel Neisseria
CC   meningitidis mutant polypeptides of the surface antigen NhhA
CC   (AAU06182-AAU06186). The modified or mutant NhhA polypeptides are
CC   characterised by deletions of non-conserved amino acids, particularly
CC   the deletion of variable regions. The deletion mutants are useful in
CC   diagnostics, therapeutic and prophylactic vaccines against a broader
CC   spectrum of N. meningitidis, and in designing and/or screening of
CC   medicaments. The mutant proteins when used as a vaccine can effectively
CC   immunise against a broader spectrum of N. meningitidis strains than
CC   would be expected from a corresponding wild-type surface antigen.
CC   The present sequence represents N. meningitidis strain PMC21 surface
CC   antigen NhhA deletion mutant #3.
XX
SQ   Sequence   433 AA;

  Query Match             100.0%;  Score 1947;  DB 22;  Length 433;
  Best Local Similarity   100.0%;  Pred. No. 6.5e-122;
  Matches  382;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 NRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKK 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         52 NRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKK 111

Qy         61 TEVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        112 TEVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTA 171

Qy        121 NGQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        172 NGQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSG 231

Qy        181 WNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSS 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        232 WNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSS 291

Qy        241 VSLGAGADAPTLSVDGDALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNR 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        292 VSLGAGADAPTLSVDGDALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNR 351

Qy        301 IDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWII 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        352 IDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWII 411

Qy        361 KGTASGNSRGHFGASASVGYQW 382
              ||||||||||||||||||||||
Db        412 KGTASGNSRGHFGASASVGYQW 433


RESULT 2
AAU06182
ID   AAU06182 standard; Protein; 512 AA.
XX
AC   AAU06182;
XX
DT   24-OCT-2001  (first entry)
XX
DE   N. meningitidis PMC21 NhhA deletion mutant #1.
XX
KW   Surface antigen NhhA; meningococcal disease; meningitis vaccine;
KW   mutant; mutein.
XX
OS   Neisseria meningitidis strain PMC21.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..51
FT                   /label= Signal_peptide
FT   Protein         52..512
FT                   /label= Mature_NhhA_deletion_mutant_#1
FT                   /note= "Predicted mature protein, specifically
FT                         claimed in claim 12"
XX
PN   WO200155182-A1.
XX
PD   02-AUG-2001.
XX
PF   25-JAN-2001; 2001WO-AU00069.
XX
PR   25-JAN-2000; 2000US-0177917.
XX
PA   (UYQU ) UNIV QUEENSLAND.
XX
PI   Peak IRA,  Jennings MP;
XX
DR   WPI; 2001-488774/53.
DR   N-PSDB; AAS09172.
XX
PT   New NhhA surface antigen polypeptides and polynucleotides from
PT   Neisseria meningitidis, useful in producing vaccines for treating or
PT   preventing broad spectrum of Neisseria meningitidis -
XX
PS   Claim 12; Fig 5; 91pp; English.
XX
CC   The present invention relates to the isolation of novel Neisseria
CC   meningitidis mutant polypeptides of the surface antigen NhhA
CC   (AAU06182-AAU06186). The modified or mutant NhhA polypeptides are
CC   characterised by deletions of non-conserved amino acids, particularly
CC   the deletion of variable regions. The deletion mutants are useful in
CC   diagnostics, therapeutic and prophylactic vaccines against a broader
CC   spectrum of N. meningitidis, and in designing and/or screening of
CC   medicaments. The mutant proteins when used as a vaccine can effectively
CC   immunise against a broader spectrum of N. meningitidis strains than
CC   would be expected from a corresponding wild-type surface antigen.
CC   The present sequence represents N. meningitidis strain PMC21 surface
CC   antigen NhhA deletion mutant #1.
XX
SQ   Sequence   512 AA;

  Query Match             99.7%;  Score 1941;  DB 22;  Length 512;
  Best Local Similarity   100.0%;  Pred. No. 2e-121;
  Matches  381;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKT 61
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        132 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKT 191

Qy         62 EVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        192 EVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 251

Qy        122 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGW 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        252 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGW 311

Qy        182 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        312 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 371

Qy        242 SLGAGADAPTLSVDGDALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRI 301
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        372 SLGAGADAPTLSVDGDALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRI 431

Qy        302 DNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIK 361
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        432 DNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIK 491

Qy        362 GTASGNSRGHFGASASVGYQW 382
                ||||||||||||||||||||
Db        492 GTASGNSRGHFGASASVGYQW 512


RESULT 3
AAY27202
ID   AAY27202 standard; Protein; 591 AA.
XX
AC   AAY27202;
XX
DT   24-SEP-1999  (first entry)
XX
DE   Amino acid sequence of N. meningitidis protein ORF40-1.
XX
KW   Neisseria meningitidis protein; pharmaceutical; vaccine; diagnosis;
KW   bacterial infection; treatment.
XX
OS   Neisseria meningitidis.
XX
PN   WO9936544-A2.
XX
PD   22-JUL-1999.
XX
PF   14-JAN-1999;   99WO-IB00103.
XX
PR   09-OCT-1998;   98GB-0022143.
PR   14-JAN-1998;   98GB-0000760.
PR   01-SEP-1998;   98GB-0019015.
XX
PA   (CHIR-) CHIRON SPA.
XX
PI   Grandi G,  Masignani V,  Pizza M,  Rappuoli R,  Scarlato V;
XX
DR   WPI; 1999-444400/37.
DR   N-PSDB; AAX99124.
XX
PT   New protein and its nucleotide sequence, useful in vaccines or
PT   diagnostic compositions for treating and/or preventing Neisseria
PT   meningitidis infections
XX
PS   Claim 1; Page 62; 123pp; English.
XX
CC   The invention provides proteins (AAY27201-245) from Neisseria
CC   meningitidis (strains A and B) and nucleic acid sequences (AAX99123-167)
CC   encoding the proteins. Compositions comprising the protein, nucleic acid
CC   or antibody specific to the protein are useful as pharmaceuticals, e.g. a
CC   vaccine composition or a diagnostic composition. The composition is also
CC   useful for treating or preventing an infection due to Neisserial
CC   bacteria, especially Neisseria meningitidis.
XX
SQ   Sequence   591 AA;

  Query Match             99.7%;  Score 1941;  DB 20;  Length 591;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-121;
  Matches  381;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKT 61
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        211 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKT 270

Qy         62 EVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        271 EVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 330

Qy        122 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGW 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        331 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGW 390

Qy        182 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 241
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        391 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 450

Qy        242 SLGAGADAPTLSVDGDALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRI 301
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        451 SLGAGADAPTLSVDGDALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRI 510

Qy        302 DNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIK 361
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        511 DNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIK 570

Qy        362 GTASGNSRGHFGASASVGYQW 382
              ||||||||||||||||||||
Db        571 GTASGNSRGHFGASASVGYQW 591


RESULT 4
AAY23741
ID   AAY23741 standard; Protein; 591 AA.
XX
AC   AAY23741;
XX
DT   08-SEP-1999  (first entry)
XX
DE   A surface protein of Neisseria meningitidis.
XX
KW   Surface protein; surface glycoprotein; infection; vaccine;
KW   immunoreactive peptide.
XX
OS   Neisseria meningitidis.
XX
PN   WO9931132-A1.
XX
PD   24-JUN-1999.
XX
PF   14-DEC-1998;   98WO-AU01031.
XX
PR   12-DEC-1997;   97GB-0026398.
XX
PA   (ISIS-) ISIS INNOVATION LTD.
PA   (UYQU ) UNIV QUEENSLAND.
XX
PI   Jennings MP,  Moxon ER,  Peak IRA;
XX
DR   WPI; 1999-418754/35.
DR   N-PSDB; AAX85793.
XX
PT   Neisseria meningitidis surface proteins useful for treating N.
PT   meningitidis infections
XX
PS   Claim 1; Page 104-106; 132pp; English.
XX
CC   The present sequence represents a surface protein of Neiserria
CC   meningitidis which is approximately 62 kDa. The N. meningitidis
CC   surface glycoproteins, nucleic acids, the primers and optionally
CC   a thermostable polymerase, or antibodies are useful in a kit for
CC   the detection or diagnosis of N. meningitidis infection in humans.
CC   The N. meningitidis surface glycoproteins can also be used to
CC   prevent or treat N. meningitidis infection in humans, especially
CC   in the form of vaccines. The proteins and antibodies can also
CC   be used to identify immunoreactive peptides.
XX
SQ   Sequence   591 AA;

  Query Match             99.7%;  Score 1941;  DB 20;  Length 591;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-121;
  Matches  381;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKT 61
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        211 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKT 270

Qy         62 EVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 121
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        271 EVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 330

Qy        122 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGW 181
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        331 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGW 390

Qy        182 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 241
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        391 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 450

Qy        242 SLGAGADAPTLSVDGDALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRI 301
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        451 SLGAGADAPTLSVDGDALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRI 510

Qy        302 DNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIK 361
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        511 DNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIK 570

Qy        362 GTASGNSRGHFGASASVGYQW 382
                 |||||||||||||||||||
Db        571 GTASGNSRGHFGASASVGYQW 591
```

RESULT 5
AAY23746
ID AAY23746 standard; Protein; 591 AA.
XX
AC AAY23746;
XX
DT 08-SEP-1999 (first entry)
XX
DE A surface protein of Neisseria meningitidis.
XX
KW Surface protein; surface glycoprotein; infection; vaccine;
KW immunoreactive peptide.
XX
OS Neisseria meningitidis.
XX
PN WO9931132-A1.
XX
PD 24-JUN-1999.
XX
PF 14-DEC-1998; 98WO-AU01031.
XX
PR 12-DEC-1997; 97GB-0026398.
XX
PA (ISIS-) ISIS INNOVATION LTD.
PA (UYQU) UNIV QUEENSLAND.
XX
PI Jennings MP, Moxon ER, Peak IRA;
XX
DR WPI; 1999-418754/35.
DR N-PSDB; AAX85798.
XX
PT Neisseria meningitidis surface proteins useful for treating N.
PT meningitidis infections
XX
PS Claim 1; Page 127-128; 132pp; English.
XX
CC The present sequence represents a surface protein of Neiserria
CC meningitidis which is approximately 62 kDa. The N. meningitidis
CC surface glycoproteins, nucleic acids, the primers and optionally
CC a thermostable polymerase, or antibodies are useful in a kit for
CC the detection or diagnosis of N. meningitidis infection in humans.
CC The N. meningitidis surface glycoproteins can also be used to
CC prevent or treat N. meningitidis infection in humans, especially
CC in the form of vaccines. The proteins and antibodies can also
CC be used to identify immunoreactive peptides.
XX
SQ Sequence 591 AA;

Query Match 99.7%; Score 1941; DB 20; Length 591;
Best Local Similarity 100.0%; Pred. No. 2.4e-121;
Matches 381; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          2 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKT 61
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        211 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKT 270

Qy         62 EVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        271 EVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 330

Qy        122 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGW 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        331 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGW 390

Qy        182 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        391 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 450

Qy        242 SLGAGADAPTLSVDGDALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRI 301
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        451 SLGAGADAPTLSVDGDALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRI 510

Qy        302 DNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIK 361
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        511 DNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIK 570

Qy        362 GTASGNSRGHFGASASVGYQW 382
              |||||||||||||||||||||
Db        571 GTASGNSRGHFGASASVGYQW 591
```

RESULT 6
AAU06171
ID AAU06171 standard; Protein; 591 AA.
XX
AC AAU06171;
XX
DT 24-OCT-2001 (first entry)
XX
DE N. meningitidis PMC21 surface antigen NhhA polypeptide sequence.
XX
KW Surface antigen NhhA; meningococcal disease; meningitis vaccine.
XX
OS Neisseria meningitidis strain PMC21.
XX
FH Key Location/Qualifiers
FT Peptide 1..51
FT /label= Signal_peptide
FT Region 1..50
FT /label= C1
FT /note= "Conserved region 1"
FT Region 51..108
FT /label= V1
FT /note= "Variable region 1"
FT Protein 52..591
FT /label= Mature_NhhA
FT /note= "Predicted mature protein, specifically
FT claimed in claim 12"
FT Region 109..120
FT /label= C2
FT /note= "Conserved region 2"
FT Region 121..124
FT /label= V2
FT /note= "Variable region 2"
FT Region 125..188
FT /label= C3
FT /note= "Conserved region 3"
FT Region 189..210
FT /label= V3
FT /note= "Variable region 3"
FT Region 211..229
FT /label= C4

```
FT                   /note= "Conserved region 4"
FT   Region          230..236
FT                   /label= V4
FT                   /note= "Variable region 4"
FT   Region          237..591
FT                   /label= C5
FT                   /note= "Conserved region 5"
XX
PN   WO200155182-A1.
XX
PD   02-AUG-2001.
XX
PF   25-JAN-2001; 2001WO-AU00069.
XX
PR   25-JAN-2000; 2000US-0177917.
XX
PA   (UYQU ) UNIV QUEENSLAND.
XX
PI   Peak IRA,  Jennings MP;
XX
DR   WPI; 2001-488774/53.
DR   N-PSDB; AAS09161.
XX
PT   New NhhA surface antigen polypeptides and polynucleotides from
PT   Neisseria meningitidis, useful in producing vaccines for treating or
PT   preventing broad spectrum of Neisseria meningitidis -
XX
PS   Claim 9; Fig 1; 91pp; English.
XX
CC   The present invention relates to the isolation of novel Neisseria
CC   meningitidis mutant polypeptides of the surface antigen NhhA
CC   (AAU06182-AAU06186). The modified or mutant NhhA polypeptides are
CC   characterised by deletions of non-conserved amino acids, particularly
CC   the deletion of variable regions. The deletion mutants are useful in
CC   diagnostics, therapeutic and prophylactic vaccines against a broader
CC   spectrum of N. meningitidis, and in designing and/or screening of
CC   medicaments. The mutant proteins when used as a vaccine can effectively
CC   immunise against a broader spectrum of N. meningitidis strains than
CC   would be expected from a corresponding wild-type surface antigen.
CC   The present sequence representing the wild type surface antigen NhhA
CC   from N. meningitidis strain PMC21 is 1 of 10 NhhA polypeptide sequences
CC   (AAU06171-AAU06180) from 10 different N. meningitidis strains given in
CC   the present invention.
XX
SQ   Sequence   591 AA;

  Query Match             99.7%;  Score 1941;  DB 22;  Length 591;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-121;
  Matches  381;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKT 61
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        211 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKT 270

Qy         62 EVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        271 EVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 330

Qy        122 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGW 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        331 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGW 390

Qy        182 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        391 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 450

Qy        242 SLGAGADAPTLSVDGDALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRI 301
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        451 SLGAGADAPTLSVDGDALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRI 510

Qy        302 DNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIK 361
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        511 DNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIK 570

Qy        362 GTASGNSRGHFGASASVGYQW 382
              |||||||||||||||||||||
Db        571 GTASGNSRGHFGASASVGYQW 591


RESULT 7
AAU06175
ID   AAU06175 standard; Protein; 591 AA.
XX
AC   AAU06175;
XX
DT   24-OCT-2001  (first entry)
XX
DE   N. meningitidis EG329 surface antigen NhhA polypeptide sequence.
XX
KW   Surface antigen NhhA; meningococcal disease; meningitis vaccine.
XX
OS   Neisseria meningitidis strain EG329.
XX
FH   Key             Location/Qualifiers
FT   Region          1..50
FT                   /label= C1
FT                   /note= "Conserved region 1"
FT   Region          51..108
FT                   /label= V1
FT                   /note= "Variable region 1"
FT   Region          109..120
FT                   /label= C2
FT                   /note= "Conserved region 2"
FT   Region          121..124
FT                   /label= V2
FT                   /note= "Variable region 2"
FT   Region          125..188
FT                   /label= C3
FT                   /note= "Conserved region 3"
FT   Region          189..210
FT                   /label= V3
FT                   /note= "Variable region 3"
FT   Region          211..229
FT                   /label= C4
FT                   /note= "Conserved region 4"
FT   Region          230..236
FT                   /label= V4
FT                   /note= "Variable region 4"
FT   Region          237..591
FT                   /label= C5
FT                   /note= "Conserved region 5"
XX
PN   WO200155182-A1.
XX
PD   02-AUG-2001.
XX
PF   25-JAN-2001; 2001WO-AU00069.
XX
PR   25-JAN-2000; 2000US-0177917.
XX
PA   (UYQU ) UNIV QUEENSLAND.
XX
PI   Peak IRA,  Jennings MP;
XX
DR   WPI; 2001-488774/53.
DR   N-PSDB; AAS09165.
XX
PT   New NhhA surface antigen polypeptides and polynucleotides from
PT   Neisseria meningitidis, useful in producing vaccines for treating or
PT   preventing broad spectrum of Neisseria meningitidis -
XX
PS   Claim 9; Fig 1; 91pp; English.
XX
CC   The present invention relates to the isolation of novel Neisseria
CC   meningitidis mutant polypeptides of the surface antigen NhhA
CC   (AAU06182-AAU06186). The modified or mutant NhhA polypeptides are
```

```
CC   characterised by deletions of non-conserved amino acids, particularly
CC   the deletion of variable regions. The deletion mutants are useful in
CC   diagnostics, therapeutic and prophylactic vaccines against a broader
CC   spectrum of N. meningitidis, and in designing and/or screening of
CC   medicaments. The mutant proteins when used as a vaccine can effectively
CC   immunise against a broader spectrum of N. meningitidis strains than
CC   would be expected from a corresponding wild-type surface antigen.
CC   The present sequence representing the wild type surface antigen NhhA
CC   from N. meningitidis strain EG329 is 1 of 10 NhhA polypeptide sequences
CC   (AAU06171-AAU06180) from 10 different N. meningitidis strains given in
CC   the present invention.
XX
SQ   Sequence   591 AA;

  Query Match             99.7%;  Score 1941;  DB 22;  Length 591;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-121;
  Matches  381;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKT 61
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        211 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKT 270

Qy         62 EVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        271 EVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 330

Qy        122 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGW 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        331 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGW 390

Qy        182 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        391 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 450

Qy        242 SLGAGADAPTLSVDGDALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRI 301
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        451 SLGAGADAPTLSVDGDALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRI 510

Qy        302 DNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIK 361
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        511 DNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIK 570

Qy        362 GTASGNSRGHFGASASVGYQW 382
              |||||||||||||||||||||
Db        571 GTASGNSRGHFGASASVGYQW 591


RESULT 8
AAY57045
ID   AAY57045 standard; Protein; 591 AA.
XX
AC   AAY57045;
XX
DT   21-FEB-2000  (first entry)
XX
DE   BASB029 amino acid sequence from N. meningitidis strain H44/76.
XX
KW   BASB029; Nisseria meningitidis; surface fibril protein; HSF; diagnosis;
KW   infection; treatment; prevent; antibacterial drug.
XX
OS   Neisseria meningitidis.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 90
FT                   /note= "Encoded by AAT"
FT   Misc-difference 92
FT                   /note= "Encoded by GAT"
FT   Misc-difference 98
FT                   /note= "Encoded by AAC"
FT   Misc-difference 108
FT                   /note= "Encoded by AATC"
FT   Misc-difference 123
FT                   /note= "Encoded by ACA"
FT   Misc-difference 269
FT                   /note= "Encoded by AAA"
FT   Misc-difference 389
FT                   /note= "Encoded by CGT"
XX
PN   WO9958683-A2.
XX
PD   18-NOV-1999.
XX
PF   07-MAY-1999;   99WO-EP03255.
XX
PR   13-MAY-1998;   98GB-0010276.
XX
PA   (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PI   Ruelle J;
XX
DR   WPI; 2000-053103/04.
DR   N-PSDB; AAZ39865.
XX
PT   New polypeptide from neisseria meningitidis useful for diagnosis,
PT   treatment or prevention of bacterial infections in mammal  -
XX
PS   Claim 4; Fig 2; 74pp; English.
XX
CC   This is the Nisseria meningitidis BASB029 amino acid sequence from
CC   serogroup B strain H44/76. The BASB029 protein is homologous to the
CC   Haemophilus influenzae surface fibril (HSF) protein. The invention
CC   relates to BASB029 polynucleotide sequences (AAZ39864-Z39865) and
CC   polypeptide sequences (AAY57044-Y57045) and their immunogenic fragments.
CC   BASB029 polypeptides are useful in a method of diagnosing a Neisseria
CC   meningitidis infection in a mammal. Compositions containing BASB029
CC   polynucleotides and polypeptides are useful for generating an immune
CC   response in an animal. A therapeutic composition comprising an antibody
CC   directed against BAsB029 is useful in treating humans with Neisseria
CC   meningitidis disease. The polynucleotide is useful in the diagnosis of
CC   the stage of infection, type of infection, susceptibility to an
CC   infection which results from increased or decreased expression of the
CC   polynucleotide, and for therapeutic or prophylactic purposes,
CC   particularly genetic immunisation. Antibodies against BASB029
CC   polynucleotides and polypeptides are also useful for treating infections
CC   particularly bacterial infections. The protein is useful in the
CC   screening and development of antibacterial drugs. Fused recombinant
CC   protein is useful for the stimulation of the immune system of an organism
CC   receiving the protein.
XX
SQ   Sequence   591 AA;

  Query Match             99.5%;  Score 1938;  DB 21;  Length 591;
  Best Local Similarity   99.7%;  Pred. No. 3.8e-121;
  Matches  380;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKT 61
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:|
Db        211 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKRT 270

Qy         62 EVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        271 EVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 330

Qy        122 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGW 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        331 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGW 390

Qy        182 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        391 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 450

Qy        242 SLGAGADAPTLSVDGDALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRI 301
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        451 SLGAGADAPTLSVDGDALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRI 510

Qy        302 DNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIK 361
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        511 DNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIK 570

Qy        362 GTASGNSRGHFGASASVGYQW 382
                ||||||||||||||||||||
Db        571 GTASGNSRGHFGASASVGYQW 591


RESULT 9
AAY23737
ID   AAY23737 standard; Protein; 592 AA.
XX
AC   AAY23737;
XX
DT   08-SEP-1999  (first entry)
XX
DE   A surface protein of Neisseria meningitidis.
XX
KW   Surface protein; surface glycoprotein; infection; vaccine;
KW   immunoreactive peptide.
XX
OS   Neisseria meningitidis.
XX
PN   WO9931132-A1.
XX
PD   24-JUN-1999.
XX
PF   14-DEC-1998;   98WO-AU01031.
XX
PR   12-DEC-1997;   97GB-0026398.
XX
PA   (ISIS-) ISIS INNOVATION LTD.
PA   (UYQU ) UNIV QUEENSLAND.
XX
PI   Jennings MP,  Moxon ER,  Peak IRA;
XX
DR   WPI; 1999-418754/35.
DR   N-PSDB; AAX85788.
XX
PT   Neisseria meningitidis surface proteins useful for treating N.
PT   meningitidis infections
XX
PS   Claim 1; Page 86-87; 132pp; English.
XX
CC   The present sequence represents a surface protein of Neiserria
CC   meningitidis which is approximately 62 kDa. The N. meningitidis
CC   surface glycoproteins, nucleic acids, the primers and optionally
CC   a thermostable polymerase, or antibodies are useful in a kit for
CC   the detection or diagnosis of N. meningitidis infection in humans.
CC   The N. meningitidis surface glycoproteins can also be used to
CC   prevent or treat N. meningitidis infection in humans, especially
CC   in the form of vaccines. The proteins and antibodies can also
CC   be used to identify immunoreactive peptides.
XX
SQ   Sequence   592 AA;

  Query Match             99.5%;  Score 1937;  DB 20;  Length 592;
  Best Local Similarity   99.7%;  Pred. No. 4.4e-121;
  Matches  380;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKT 61
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        212 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKT 271

Qy         62 EVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 121
              ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        272 EVKIGVKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 331

Qy        122 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGW 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        332 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGW 391

Qy        182 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 241
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        392 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 451

Qy        242 SLGAGADAPTLSVDGDALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRI 301
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        452 SLGAGADAPTLSVDGDALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRI 511

Qy        302 DNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIK 361
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        512 DNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIK 571

Qy        362 GTASGNSRGHFGASASVGYQW 382
                ||||||||||||||||||||
Db        572 GTASGNSRGHFGASASVGYQW 592


RESULT 10
AAU06186
ID   AAU06186 standard; Protein; 502 AA.
XX
AC   AAU06186;
XX
DT   24-OCT-2001  (first entry)
XX
DE   N. meningitidis PMC21 NhhA deletion mutant #4.
XX
KW   Surface antigen NhhA; meningococcal disease; meningitis vaccine;
KW   mutant; mutein.
XX
OS   Neisseria meningitidis strain PMC21.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..49
FT                   /label= Signal_peptide
FT   Protein         50..502
FT                   /label= Mature_NhhA_deletion_mutant_#4
FT                   /note= "Predicted mature protein, specifically
FT                         claimed in claim 12"
XX
PN   WO200155182-A1.
XX
PD   02-AUG-2001.
XX
PF   25-JAN-2001; 2001WO-AU00069.
XX
PR   25-JAN-2000; 2000US-0177917.
XX
PA   (UYQU ) UNIV QUEENSLAND.
XX
PI   Peak IRA,  Jennings MP;
XX
DR   WPI; 2001-488774/53.
DR   N-PSDB; AAS09176.
XX
PT   New NhhA surface antigen polypeptides and polynucleotides from
PT   Neisseria meningitidis, useful in producing vaccines for treating or
PT   preventing broad spectrum of Neisseria meningitidis -
XX
PS   Claim 12; Fig 9; 91pp; English.
XX
CC   The present invention relates to the isolation of novel Neisseria
CC   meningitidis mutant polypeptides of the surface antigen NhhA
CC   (AAU06182-AAU06186). The modified or mutant NhhA polypeptides are
CC   characterised by deletions of non-conserved amino acids, particularly
CC   the deletion of variable regions. The deletion mutants are useful in
CC   diagnostics, therapeutic and prophylactic vaccines against a broader
CC   spectrum of N. meningitidis, and in designing and/or screening of
CC   medicaments. The mutant proteins when used as a vaccine can effectively
CC   immunise against a broader spectrum of N. meningitidis strains than
CC   would be expected from a corresponding wild-type surface antigen.
CC   The present sequence represents N. meningitidis strain PMC21 surface
CC   antigen NhhA deletion mutant #4.
```

```
XX
SQ   Sequence   502 AA;

  Query Match            97.1%;  Score 1891.5;  DB 22;  Length 502;
  Best Local Similarity  97.9%;  Pred. No. 3.9e-118;
  Matches  374;  Conservative    1;  Mismatches    0;  Indels    7;  Gaps    1;

Qy          1 NRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKK 60
              :|||||||||||||||||||       |||||||||||||||||||||||||||||||||
Db        128 DRAASVKDVLNAGWNIKGVK-------NVDFVRTYDTVEFLSADTKTTTVNVESKDNGKK 180

Qy         61 TEVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTA 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TEVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTA 240

Qy        121 NGQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSG 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 NGQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSG 300

Qy        181 WNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSS 240
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 WNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSS 360

Qy        241 VSLGAGADAPTLSVDGDALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNR 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 VSLGAGADAPTLSVDGDALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNR 420

Qy        301 IDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWII 360
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 IDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWII 480

Qy        361 KGTASGNSRGHFGASASVGYQW 382
              ||||||||||||||||||||||
Db        481 KGTASGNSRGHFGASASVGYQW 502


RESULT 11
AAY23740
ID   AAY23740 standard; Protein; 594 AA.
XX
AC   AAY23740;
XX
DT   08-SEP-1999  (first entry)
XX
DE   A surface protein of Neisseria meningitidis.
XX
KW   Surface protein; surface glycoprotein; infection; vaccine;
KW   immunoreactive peptide.
XX
OS   Neisseria meningitidis.
XX
PN   WO9931132-A1.
XX
PD   24-JUN-1999.
XX
PF   14-DEC-1998;   98WO-AU01031.
XX
PR   12-DEC-1997;   97GB-0026398.
XX
PA   (ISIS-) ISIS INNOVATION LTD.
PA   (UYQU ) UNIV QUEENSLAND.
XX
PI   Jennings MP,  Moxon ER,  Peak IRA;
XX
DR   WPI; 1999-418754/35.
DR   N-PSDB; AAX85792.
XX
PT   Neisseria meningitidis surface proteins useful for treating N.
PT   meningitidis infections
XX
PS   Claim 1; Page 100-101; 132pp; English.
XX
CC   The present sequence represents a surface protein of Neiserria
CC   meningitidis which is approximately 62 kDa. The N. meningitidis
CC   surface glycoproteins, nucleic acids, the primers and optionally
CC   a thermostable polymerase, or antibodies are useful in a kit for
CC   the detection or diagnosis of N. meningitidis infection in humans.
CC   The N. meningitidis surface glycoproteins can also be used to
CC   prevent or treat N. meningitidis infection in humans, especially
CC   in the form of vaccines. The proteins and antibodies can also
CC   be used to identify immunoreactive peptides.
XX
SQ   Sequence   594 AA;

  Query Match            96.9%;  Score 1886.5;  DB 20;  Length 594;
  Best Local Similarity  97.6%;  Pred. No. 1e-117;
  Matches  373;  Conservative    3;  Mismatches    5;  Indels    1;  Gaps    1;

Qy          2 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKT 61
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:|
Db        213 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKRT 272

Qy         62 EVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 121
              |||||||||||||||||||||||| |||:|||||||||||||||||||||||||||||||
Db        273 EVKIGAKTSVIKEKDGKLVTGKDKGENDSSTDKGEGLVTAKEVIDAVNKAGWRMKTTTAN 332

Qy        122 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGW 181
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        333 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGW 392

Qy        182 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 241
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        393 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 452

Qy        242 SLGAGADAPTLSVDGD-ALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNR 300
              |||||||||||||| : ||||||| |||||||||||||||||||||||||||||||||||
Db        453 SLGAGADAPTLSVDDEGALNVGSKDANKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNH 512

Qy        301 IDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWII 360
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        513 IDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWII 572

Qy        361 KGTASGNSRGHFGASASVGYQW 382
              ||||||||||||||||||||||
Db        573 KGTASGNSRGHFGASASVGYQW 594


RESULT 12
AAY57044
ID   AAY57044 standard; Protein; 594 AA.
XX
AC   AAY57044;
XX
DT   21-FEB-2000  (first entry)
XX
DE   BASB029 amino acid sequence from N. meningitidis strain ATCC13090.
XX
KW   BASB029; Nisseria meningitidis; surface fibril protein; HSF; diagnosis;
KW   infection; treatment; prevent; antibacterial drug.
XX
OS   Neisseria meningitidis.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 104
FT                   /note= "Encoded by AATC"
XX
PN   WO9958683-A2.
XX
PD   18-NOV-1999.
XX
PF   07-MAY-1999;   99WO-EP03255.
XX
PR   13-MAY-1998;   98GB-0010276.
XX
PA   (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX
```

```
PI   Ruelle J;
XX
DR   WPI; 2000-053103/04.
DR   N-PSDB; AAZ39864.
XX
PT   New polypeptide from neisseria meningitidis useful for diagnosis,
PT   treatment or prevention of bacterial infections in mammal -
XX
PS   Claim 4; Fig 2; 74pp; English.
XX
CC   This is the Nisseria meningitidis BASB029 amino acid sequence from
CC   serogroup B strain ATCC13090. The BASB029 protein is homologous to the
CC   Haemophilus influenzae surface fibril (HSF) protein. The invention
CC   relates to BASB029 polynucleotide sequences (AAZ39864-Z39865) and
CC   polypeptide sequences (AAY57044-Y57045) and their immunogenic fragments.
CC   BASB029 polypeptides are useful in a method of diagnosing a Neisseria
CC   meningitidis infection in a mammal. Compositions containing BASB029
CC   polynucleotides and polypeptides are useful for generating an immune
CC   response in an animal. A therapeutic composition comprising an antibody
CC   directed against BAsB029 is useful in treating humans with Neisseria
CC   meningitidis disease. The polynucleotide is useful in the diagnosis of
CC   the stage of infection, type of infection, susceptibility to an
CC   infection which results from increased or decreased expression of the
CC   polynucleotide, and for therapeutic or prophylactic purposes,
CC   particularly genetic immunisation. Antibodies against BASB029
CC   polynucleotides and polypeptides are also useful for treating infections
CC   particularly bacterial infections. The protein is useful in the
CC   screening and development of antibacterial drugs. Fused recombinant
CC   protein is useful for the stimulation of the immune system of an organism
CC   receiving the protein.
XX
SQ   Sequence   594 AA;

  Query Match            96.9%;  Score 1886.5;  DB 21;  Length 594;
  Best Local Similarity  97.6%;  Pred. No. 1e-117;
  Matches  373;  Conservative    3;  Mismatches    5;  Indels    1;  Gaps    1;

Qy          2 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKT 61
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:|
Db        213 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKRT 272

Qy         62 EVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 121
              |||||||||||||||||||||||||| |||||:|||||||||||||||||||||||||||
Db        273 EVKIGAKTSVIKEKDGKLVTGKDKGENDSSTDKGEGLVTAKEVIDAVNKAGWRMKTTTAN 332

Qy        122 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGW 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        333 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGW 392

Qy        182 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        393 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 452

Qy        242 SLGAGADAPTLSVDGD-ALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNR 300
              |||||||||||||| : |||||||  ||||||||||||||||||||||||||||||||||
Db        453 SLGAGADAPTLSVDDEGALNVGSKDANKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNH 512

Qy        301 IDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWII 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        513 IDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWII 572

Qy        361 KGTASGNSRGHFGASASVGYQW 382
              ||||||||||||||||||||||
Db        573 KGTASGNSRGHFGASASVGYQW 594


RESULT 13
AAU06174
ID   AAU06174 standard; Protein; 594 AA.
XX
AC   AAU06174;
XX
DT   24-OCT-2001  (first entry)
XX
DE   N. meningitidis EG327 surface antigen NhhA polypeptide sequence.
XX
KW   Surface antigen NhhA; meningococcal disease; meningitis vaccine.
XX
OS   Neisseria meningitidis strain EG327.
XX
FH   Key             Location/Qualifiers
FT   Region          1..50
FT                   /label= C1
FT                   /note= "Conserved region 1"
FT   Region          51..104
FT                   /label= V1
FT                   /note= "Variable region 1"
FT   Region          105..116
FT                   /label= C2
FT                   /note= "Conserved region 2"
FT   Region          117..126
FT                   /label= V2
FT                   /note= "Variable region 2"
FT   Region          127..190
FT                   /label= C3
FT                   /note= "Conserved region 3"
FT   Region          191..212
FT                   /label= V3
FT                   /note= "Variable region 3"
FT   Region          213..231
FT                   /label= C4
FT                   /note= "Conserved region 4"
FT   Region          232..238
FT                   /label= V4
FT                   /note= "Variable region 4"
FT   Region          239..594
FT                   /label= C5
FT                   /note= "Conserved region 5"
XX
PN   WO200155182-A1.
XX
PD   02-AUG-2001.
XX
PF   25-JAN-2001; 2001WO-AU00069.
XX
PR   25-JAN-2000; 2000US-0177917.
XX
PA   (UYQU ) UNIV QUEENSLAND.
XX
PI   Peak IRA,  Jennings MP;
XX
DR   WPI; 2001-488774/53.
DR   N-PSDB; AAS09164.
XX
PT   New NhhA surface antigen polypeptides and polynucleotides from
PT   Neisseria meningitidis, useful in producing vaccines for treating or
PT   preventing broad spectrum of Neisseria meningitidis -
XX
PS   Claim 9; Fig 1; 91pp; English.
XX
CC   The present invention relates to the isolation of novel Neisseria
CC   meningitidis mutant polypeptides of the surface antigen NhhA
CC   (AAU06182-AAU06186). The modified or mutant NhhA polypeptides are
CC   characterised by deletions of non-conserved amino acids, particularly
CC   the deletion of variable regions. The deletion mutants are useful in
CC   diagnostics, therapeutic and prophylactic vaccines against a broader
CC   spectrum of N. meningitidis, and in designing and/or screening of
CC   medicaments. The mutant proteins when used as a vaccine can effectively
CC   immunise against a broader spectrum of N. meningitidis strains than
CC   would be expected from a corresponding wild-type surface antigen.
CC   The present sequence representing the wild type surface antigen NhhA
CC   from N. meningitidis strain EG327 is 1 of 10 NhhA polypeptide sequences
CC   (AAU06171-AAU06180) from 10 different N. meningitidis strains given in
CC   the present invention.
XX
SQ   Sequence   594 AA;
```

```
  Query Match             96.9%;  Score 1886.5;  DB 22;  Length 594;
  Best Local Similarity   97.6%;  Pred. No. 1e-117;
  Matches  373;  Conservative    3;  Mismatches    5;  Indels    1;  Gaps    1;

Qy          2 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKT 61
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:|
Db        213 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKRT 272

Qy         62 EVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 121
              |||||||||||||||||||||||||| |||||:|||||||||||||||||||||||||||
Db        273 EVKIGAKTSVIKEKDGKLVTGKDKGENDSSTDKGEGLVTAKEVIDAVNKAGWRMKTTTAN 332

Qy        122 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGW 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        333 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGW 392

Qy        182 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        393 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 452

Qy        242 SLGAGADAPTLSVDGD-ALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNR 300
              ||||||||||||||  : |||||||  ||||||||||||||||||||||||||||||||
Db        453 SLGAGADAPTLSVDDEGALNVGSKDANKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNH 512

Qy        301 IDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWII 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        513 IDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWII 572

Qy        361 KGTASGNSRGHFGASASVGYQW 382
              ||||||||||||||||||||||
Db        573 KGTASGNSRGHFGASASVGYQW 594


RESULT 14
AAY23743
ID   AAY23743 standard; Protein; 599 AA.
XX
AC   AAY23743;
XX
DT   08-SEP-1999  (first entry)
XX
DE   A surface protein of Neisseria meningitidis.
XX
KW   Surface protein; surface glycoprotein; infection; vaccine;
KW   immunoreactive peptide.
XX
OS   Neisseria meningitidis.
XX
PN   WO9931132-A1.
XX
PD   24-JUN-1999.
XX
PF   14-DEC-1998;   98WO-AU01031.
XX
PR   12-DEC-1997;   97GB-0026398.
XX
PA   (ISIS-) ISIS INNOVATION LTD.
PA   (UYQU ) UNIV QUEENSLAND.
XX
PI   Jennings MP,  Moxon ER,  Peak IRA;
XX
DR   WPI; 1999-418754/35.
DR   N-PSDB; AAX85795.
XX
PT   Neisseria meningitidis surface proteins useful for treating N.
PT   meningitidis infections
XX
PS   Claim 1; Page 114-115; 132pp; English.
XX
CC   The present sequence represents a surface protein of Neiserria
CC   meningitidis which is approximately 62 kDa. The N. meningitidis
CC   surface glycoproteins, nucleic acids, the primers and optionally
CC   a thermostable polymerase, or antibodies are useful in a kit for
CC   the detection or diagnosis of N. meningitidis infection in humans.
CC   The N. meningitidis surface glycoproteins can also be used to
CC   prevent or treat N. meningitidis infection in humans, especially
CC   in the form of vaccines. The proteins and antibodies can also
CC   be used to identify immunoreactive peptides.
XX
SQ   Sequence   599 AA;

  Query Match             96.6%;  Score 1881.5;  DB 20;  Length 599;
  Best Local Similarity   97.6%;  Pred. No. 2.3e-117;
  Matches  373;  Conservative    1;  Mismatches    7;  Indels    1;  Gaps    1;

Qy          2 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKT 61
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||:|
Db        218 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVHTYDTVEFLSADTKTTTVNVESKDNGKRT 277

Qy         62 EVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 121
              |||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db        278 EVKIGAKTSVIKEKDGKLVTGKGKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 337

Qy        122 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGW 181
              |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        338 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVKYDVNVGDALNVNQLQNSGW 397

Qy        182 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        398 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 457

Qy        242 SLGAGADAPTLSVDG-DALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNR 300
              |||||||||||||   ||||||| ||||||||||||||||||||||||||||||||||||
Db        458 SLGAGADAPTLSVDDKGALNVGSKDANKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNR 517

Qy        301 IDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWII 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        518 IDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWII 577

Qy        361 KGTASGNSRGHFGASASVGYQW 382
              ||||||||||||||||||||||
Db        578 KGTASGNSRGHFGASASVGYQW 599


RESULT 15
AAU06176
ID   AAU06176 standard; Protein; 599 AA.
XX
AC   AAU06176;
XX
DT   24-OCT-2001  (first entry)
XX
DE   N. meningitidis H38 surface antigen NhhA polypeptide sequence.
XX
KW   Surface antigen NhhA; meningococcal disease; meningitis vaccine.
XX
OS   Neisseria meningitidis strain H38.
XX
FH   Key             Location/Qualifiers
FT   Region          1..50
FT                   /label= C1
FT                   /note= "Conserved region 1"
FT   Region          51..105
FT                   /label= V1
FT                   /note= "Variable region 1"
FT   Region          106..117
FT                   /label= C2
FT                   /note= "Conserved region 2"
FT   Region          118..131
FT                   /label= V2
FT                   /note= "Variable region 2"
FT   Region          132..195
FT                   /label= C3
FT                   /note= "Conserved region 3"
FT   Region          196..217
```

```
FT                   /label= V3
FT                   /note= "Variable region 3"
FT   Region          218..236
FT                   /label= C4
FT                   /note= "Conserved region 4"
FT   Region          237..243
FT                   /label= V4
FT                   /note= "Variable region 4"
FT   Region          244..599
FT                   /label= C5
FT                   /note= "Conserved region 5"
XX
PN   WO200155182-A1.
XX
PD   02-AUG-2001.
XX
PF   25-JAN-2001; 2001WO-AU00069.
XX
PR   25-JAN-2000; 2000US-0177917.
XX
PA   (UYQU ) UNIV QUEENSLAND.
XX
PI   Peak IRA,  Jennings MP;
XX
DR   WPI; 2001-488774/53.
DR   N-PSDB; AAS09166.
XX
PT   New NhhA surface antigen polypeptides and polynucleotides from
PT   Neisseria meningitidis, useful in producing vaccines for treating or
PT   preventing broad spectrum of Neisseria meningitidis -
XX
PS   Claim 9; Fig 1; 91pp; English.
XX
CC   The present invention relates to the isolation of novel Neisseria
CC   meningitidis mutant polypeptides of the surface antigen NhhA
CC   (AAU06182-AAU06186). The modified or mutant NhhA polypeptides are
CC   characterised by deletions of non-conserved amino acids, particularly
CC   the deletion of variable regions. The deletion mutants are useful in
CC   diagnostics, therapeutic and prophylactic vaccines against a broader
CC   spectrum of N. meningitidis, and in designing and/or screening of
CC   medicaments. The mutant proteins when used as a vaccine can effectively
CC   immunise against a broader spectrum of N. meningitidis strains than
CC   would be expected from a corresponding wild-type surface antigen.
CC   The present sequence representing the wild type surface antigen NhhA
CC   from N. meningitidis strain H38 is 1 of 10 NhhA polypeptide sequences
CC   (AAU06171-AAU06180) from 10 different N. meningitidis strains given in
CC   the present invention.
XX
SQ   Sequence   599 AA;

  Query Match            96.6%;  Score 1881.5;  DB 22;  Length 599;
  Best Local Similarity  97.6%;  Pred. No. 2.3e-117;
  Matches  373;  Conservative    1;  Mismatches    7;  Indels    1;  Gaps    1;

Qy          2 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKT 61
              |||||||||||||||||||||||||||||| |||||||||||||||||||||||||:|
Db        218 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVHTYDTVEFLSADTKTTTVNVESKDNGKRT 277

Qy         62 EVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 121
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        278 EVKIGAKTSVIKEKDGKLVTGKGKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 337

Qy        122 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGW 181
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        338 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVKYDVNVGDALNVNQLQNSGW 397

Qy        182 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 241
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        398 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 457

Qy        242 SLGAGADAPTLSVDG-DALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNR 300
              |||||||||||||    |||||||  ||||||||||||||||||||||||||||||||||
Db        458 SLGAGADAPTLSVDDKGALNVGSKDANKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNR 517

Qy        301 IDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWII 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        518 IDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWII 577

Qy        361 KGTASGNSRGHFGASASVGYQW 382
              ||||||||||||||||||||||
Db        578 KGTASGNSRGHFGASASVGYQW 599


Search completed: October  6, 2003, 09:22:45
Job time : 34.8045 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: October 6, 2003, 09:14:45 ; Search time 10.2154 Seconds
(without alignments)
1582.188 Million cell updates/sec

Title: US-09-771-382-38
Perfect score: 1947
Sequence: 1 NRAASVKDVLNAGWNIKGVK.....TASGNSRGHFGASASVGYQW 382

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1941	99.7	591	3	US-09-377-155-11	Sequence 11, Appl
2	1941	99.7	591	3	US-09-377-155-21	Sequence 21, Appl
3	1941	99.7	591	4	US-09-669-974-11	Sequence 11, Appl
4	1941	99.7	591	4	US-09-669-974-21	Sequence 21, Appl
5	1937	99.5	592	3	US-09-377-155-2	Sequence 2, Appli
6	1937	99.5	592	4	US-09-669-974-2	Sequence 2, Appli
7	1886.5	96.9	594	3	US-09-377-155-9	Sequence 9, Appli
8	1886.5	96.9	594	4	US-09-669-974-9	Sequence 9, Appli
9	1881.5	96.6	599	3	US-09-377-155-15	Sequence 15, Appl
10	1881.5	96.6	599	4	US-09-669-974-15	Sequence 15, Appl
11	1874.5	96.3	594	3	US-09-377-155-7	Sequence 7, Appli
12	1874.5	96.3	594	4	US-09-669-974-7	Sequence 7, Appli
13	1861.5	95.6	592	3	US-09-377-155-17	Sequence 17, Appl
14	1861.5	95.6	592	4	US-09-669-974-17	Sequence 17, Appl
15	1861.5	95.6	598	3	US-09-377-155-5	Sequence 5, Appli
16	1861.5	95.6	598	3	US-09-377-155-13	Sequence 13, Appl
17	1861.5	95.6	598	4	US-09-669-974-5	Sequence 5, Appli
18	1861.5	95.6	598	4	US-09-669-974-13	Sequence 13, Appl
19	1817.5	93.3	589	3	US-09-377-155-19	Sequence 19, Appl
20	1817.5	93.3	589	4	US-09-669-974-19	Sequence 19, Appl
21	710	36.5	1098	1	US-08-409-995-2	Sequence 2, Appli
22	710	36.5	1098	3	US-08-685-467-2	Sequence 2, Appli
23	710	36.5	1098	3	US-09-377-155-32	Sequence 32, Appl
24	710	36.5	1098	3	US-08-913-942-2	Sequence 2, Appli
25	710	36.5	1098	4	US-09-669-974-32	Sequence 32, Appl
26	710	36.5	1098	4	US-09-268-347-44	Sequence 44, Appl
27	692	35.5	658	1	US-08-409-995-5	Sequence 5, Appli
28	692	35.5	658	3	US-08-685-467-5	Sequence 5, Appli
29	692	35.5	658	3	US-08-913-942-5	Sequence 5, Appli
30	690.5	35.5	1094	4	US-09-268-347-32	Sequence 32, Appl
31	660	33.9	2353	3	US-09-377-155-33	Sequence 33, Appl
32	660	33.9	2353	3	US-08-913-942-4	Sequence 4, Appli
33	660	33.9	2353	4	US-09-669-974-33	Sequence 33, Appl
34	660	33.9	2411	4	US-09-268-347-36	Sequence 36, Appl
35	659.5	33.9	2354	4	US-09-268-347-47	Sequence 47, Appl
36	635	32.6	1912	1	US-08-409-995-4	Sequence 4, Appli
37	635	32.6	1912	3	US-08-685-467-4	Sequence 4, Appli
38	612.5	31.5	607	1	US-08-409-995-6	Sequence 6, Appli
39	612.5	31.5	607	3	US-08-685-467-6	Sequence 6, Appli
40	612.5	31.5	607	3	US-08-913-942-6	Sequence 6, Appli
41	470.5	24.2	679	3	US-08-913-942-15	Sequence 15, Appl
42	470.5	24.2	679	4	US-09-268-347-26	Sequence 26, Appl
43	414.5	21.3	1002	4	US-09-268-347-24	Sequence 24, Appl
44	406	20.9	1004	4	US-09-268-347-30	Sequence 30, Appl
45	392	20.1	1104	4	US-09-268-347-28	Sequence 28, Appl

ALIGNMENTS

RESULT 1
US-09-377-155-11
; Sequence 11, Application US/09377155
; Patent No. 6197312
; GENERAL INFORMATION:
; APPLICANT: PEAK, Ian Richard Anselm
; APPLICANT: JENNINGS, Michael Paul
; APPLICANT: MOXON, E. Richard
; TITLE OF INVENTION: NOVEL SURFACE ANTIGEN
; FILE REFERENCE: 065064/0128
; CURRENT APPLICATION NUMBER: US/09/377,155
; CURRENT FILING DATE: 1999-08-19
; PRIOR APPLICATION NUMBER: PCT/AU98/01031
; PRIOR FILING DATE: 1998-12-14
; PRIOR APPLICATION NUMBER: GB 9726398.2
; PRIOR FILING DATE: 1997-12-12
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 591
; TYPE: PRT
; ORGANISM: Neisseria meningitidis
US-09-377-155-11

Query Match 99.7%; Score 1941; DB 3; Length 591;
Best Local Similarity 100.0%; Pred. No. 1.8e-150;
Matches 381; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          2 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKT 61
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        211 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKT 270

Qy         62 EVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        271 EVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 330

Qy        122 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGW 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        331 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGW 390

Qy        182 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        391 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 450

Qy        242 SLGAGADAPTLSVDGDALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRI 301
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        451 SLGAGADAPTLSVDGDALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRI 510

Qy        302 DNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIK 361
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        511 DNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIK 570

Qy        362 GTASGNSRGHFGASASVGYQW 382
              |||||||||||||||||||||
Db        571 GTASGNSRGHFGASASVGYQW 591


RESULT 2
US-09-377-155-21
; Sequence 21, Application US/09377155
; Patent No. 6197312
; GENERAL INFORMATION:
;  APPLICANT:  PEAK, Ian Richard Anselm
;  APPLICANT:  JENNINGS, Michael Paul
;  APPLICANT:  MOXON, E. Richard
;  TITLE OF INVENTION:  NOVEL SURFACE ANTIGEN
;  FILE REFERENCE:  065064/0128
;  CURRENT APPLICATION NUMBER: US/09/377,155
;  CURRENT FILING DATE:  1999-08-19
;  PRIOR APPLICATION NUMBER:  PCT/AU98/01031
;  PRIOR FILING DATE:  1998-12-14
;  PRIOR APPLICATION NUMBER:  GB 9726398.2
;  PRIOR FILING DATE:  1997-12-12
;  NUMBER OF SEQ ID NOS:  33
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 21
;   LENGTH: 591
;   TYPE: PRT
;   ORGANISM: Neisseria meningitidis
US-09-377-155-21

  Query Match             99.7%;  Score 1941;  DB 3;  Length 591;
  Best Local Similarity   100.0%;  Pred. No. 1.8e-150;
  Matches  381;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKT 61
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        211 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKT 270

Qy         62 EVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 121
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        271 EVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 330

Qy        122 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGW 181
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        331 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGW 390

Qy        182 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 241
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        391 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 450

Qy        242 SLGAGADAPTLSVDGDALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRI 301
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        451 SLGAGADAPTLSVDGDALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRI 510

Qy        302 DNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIK 361
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        511 DNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIK 570

Qy        362 GTASGNSRGHFGASASVGYQW 382
              |||||||||||||||||||||
Db        571 GTASGNSRGHFGASASVGYQW 591


RESULT 3
US-09-669-974-11
; Sequence 11, Application US/09669974
; Patent No. 6333173
; GENERAL INFORMATION:
;  APPLICANT: PEAK, Ian Richard Anselm
;  APPLICANT:  JENNINGS, Michael Paul
;  APPLICANT:  MOXON, E. Richard
;  TITLE OF INVENTION: NOVEL SURFACE ANTIGEN
;  FILE REFERENCE: 065064/0128
;  CURRENT APPLICATION NUMBER: US/09/669,974
;  CURRENT FILING DATE:  2000-09-26
;  PRIOR APPLICATION NUMBER: US 09/377,155
;  PRIOR FILING DATE: 1999-08-19
;  PRIOR APPLICATION NUMBER: PCT/AU98/01031
;  PRIOR FILING DATE: 1998-12-14
;  PRIOR APPLICATION NUMBER: GB 9726398.2
;  PRIOR FILING DATE: 1997-12-12
;  NUMBER OF SEQ ID NOS: 33
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
;   LENGTH: 591
;   TYPE: PRT
;   ORGANISM: Neisseria meningitidis
US-09-669-974-11

  Query Match             99.7%;  Score 1941;  DB 4;  Length 591;
  Best Local Similarity   100.0%;  Pred. No. 1.8e-150;
  Matches  381;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKT 61
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        211 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKT 270

Qy         62 EVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 121
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        271 EVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 330

Qy        122 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGW 181
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        331 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGW 390

Qy        182 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 241
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        391 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 450

Qy        242 SLGAGADAPTLSVDGDALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRI 301
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        451 SLGAGADAPTLSVDGDALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRI 510

Qy        302 DNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIK 361
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        511 DNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIK 570

Qy        362 GTASGNSRGHFGASASVGYQW 382
              |||||||||||||||||||||
Db        571 GTASGNSRGHFGASASVGYQW 591


RESULT 4
US-09-669-974-21
; Sequence 21, Application US/09669974
; Patent No. 6333173
; GENERAL INFORMATION:
;  APPLICANT: PEAK, Ian Richard Anselm
;  APPLICANT:  JENNINGS, Michael Paul
;  APPLICANT:  MOXON, E. Richard
;  TITLE OF INVENTION: NOVEL SURFACE ANTIGEN
;  FILE REFERENCE: 065064/0128
;  CURRENT APPLICATION NUMBER: US/09/669,974
;  CURRENT FILING DATE:  2000-09-26
;  PRIOR APPLICATION NUMBER: US 09/377,155
;  PRIOR FILING DATE: 1999-08-19
;  PRIOR APPLICATION NUMBER: PCT/AU98/01031
;  PRIOR FILING DATE: 1998-12-14
;  PRIOR APPLICATION NUMBER: GB 9726398.2
;  PRIOR FILING DATE: 1997-12-12
;  NUMBER OF SEQ ID NOS: 33
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 21
;   LENGTH: 591
;   TYPE: PRT
```

```
;   ORGANISM: Neisseria meningitidis
US-09-669-974-21

  Query Match            99.7%;  Score 1941;  DB 4;  Length 591;
  Best Local Similarity  100.0%;  Pred. No. 1.8e-150;
  Matches  381;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKT 61
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        211 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKT 270

Qy         62 EVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        271 EVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 330

Qy        122 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGW 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        331 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGW 390

Qy        182 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        391 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 450

Qy        242 SLGAGADAPTLSVDGDALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRI 301
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        451 SLGAGADAPTLSVDGDALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRI 510

Qy        302 DNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIK 361
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        511 DNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIK 570

Qy        362 GTASGNSRGHFGASASVGYQW 382
              |||||||||||||||||||||
Db        571 GTASGNSRGHFGASASVGYQW 591


RESULT 5
US-09-377-155-2
; Sequence 2, Application US/09377155
; Patent No. 6197312
; GENERAL INFORMATION:
;  APPLICANT:  PEAK, Ian Richard Anselm
;  APPLICANT:  JENNINGS, Michael Paul
;  APPLICANT:  MOXON, E. Richard
;  TITLE OF INVENTION:  NOVEL SURFACE ANTIGEN
;  FILE REFERENCE:  065064/0128
;  CURRENT APPLICATION NUMBER: US/09/377,155
;  CURRENT FILING DATE:  1999-08-19
;  PRIOR APPLICATION NUMBER:  PCT/AU98/01031
;  PRIOR FILING DATE:  1998-12-14
;  PRIOR APPLICATION NUMBER:  GB 9726398.2
;  PRIOR FILING DATE:  1997-12-12
;  NUMBER OF SEQ ID NOS:  33
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 592
;   TYPE: PRT
;   ORGANISM: Neisseria meningitidis
US-09-377-155-2

  Query Match            99.5%;  Score 1937;  DB 3;  Length 592;
  Best Local Similarity  99.7%;  Pred. No. 3.8e-150;
  Matches  380;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKT 61
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        212 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKT 271

Qy         62 EVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 121
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        272 EVKIGVKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 331

Qy        122 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGW 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        332 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGW 391

Qy        182 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        392 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 451

Qy        242 SLGAGADAPTLSVDGDALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRI 301
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        452 SLGAGADAPTLSVDGDALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRI 511

Qy        302 DNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIK 361
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        512 DNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIK 571

Qy        362 GTASGNSRGHFGASASVGYQW 382
              |||||||||||||||||||||
Db        572 GTASGNSRGHFGASASVGYQW 592


RESULT 6
US-09-669-974-2
; Sequence 2, Application US/09669974
; Patent No. 6333173
; GENERAL INFORMATION:
;  APPLICANT: PEAK, Ian Richard Anselm
;  APPLICANT:  JENNINGS, Michael Paul
;  APPLICANT:  MOXON, E. Richard
;  TITLE OF INVENTION: NOVEL SURFACE ANTIGEN
;  FILE REFERENCE: 065064/0128
;  CURRENT APPLICATION NUMBER: US/09/669,974
;  CURRENT FILING DATE:  2000-09-26
;  PRIOR APPLICATION NUMBER: US 09/377,155
;  PRIOR FILING DATE: 1999-08-19
;  PRIOR APPLICATION NUMBER: PCT/AU98/01031
;  PRIOR FILING DATE: 1998-12-14
;  PRIOR APPLICATION NUMBER: GB 9726398.2
;  PRIOR FILING DATE: 1997-12-12
;  NUMBER OF SEQ ID NOS: 33
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 592
;   TYPE: PRT
;   ORGANISM: Neisseria meningitidis
US-09-669-974-2

  Query Match            99.5%;  Score 1937;  DB 4;  Length 592;
  Best Local Similarity  99.7%;  Pred. No. 3.8e-150;
  Matches  380;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKT 61
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        212 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKT 271

Qy         62 EVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 121
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        272 EVKIGVKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 331

Qy        122 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGW 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        332 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGW 391

Qy        182 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        392 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 451

Qy        242 SLGAGADAPTLSVDGDALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRI 301
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        452 SLGAGADAPTLSVDGDALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRI 511

Qy        302 DNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIK 361
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        512 DNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIK 571
```

```
Qy        362 GTASGNSRGHFGASASVGYQW 382
              ||||||||||||||||||||
Db        572 GTASGNSRGHFGASASVGYQW 592


RESULT 7
US-09-377-155-9
; Sequence 9, Application US/09377155
; Patent No. 6197312
; GENERAL INFORMATION:
;  APPLICANT:  PEAK, Ian Richard Anselm
;  APPLICANT:  JENNINGS, Michael Paul
;  APPLICANT:  MOXON, E. Richard
;  TITLE OF INVENTION:  NOVEL SURFACE ANTIGEN
;  FILE REFERENCE:  065064/0128
;  CURRENT APPLICATION NUMBER: US/09/377,155
;  CURRENT FILING DATE:  1999-08-19
;  PRIOR APPLICATION NUMBER:  PCT/AU98/01031
;  PRIOR FILING DATE:  1998-12-14
;  PRIOR APPLICATION NUMBER:  GB 9726398.2
;  PRIOR FILING DATE:  1997-12-12
;  NUMBER OF SEQ ID NOS:  33
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 9
;   LENGTH: 594
;   TYPE: PRT
;   ORGANISM: Neisseria meningitidis
US-09-377-155-9

  Query Match          96.9%;  Score 1886.5;  DB 3;  Length 594;
  Best Local Similarity   97.6%;  Pred. No. 5.1e-146;
  Matches  373;  Conservative    3;  Mismatches    5;  Indels    1;  Gaps    1;

Qy          2 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKT 61
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:|
Db        213 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKRT 272

Qy         62 EVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 121
              |||||||||||||||||||||||| |||||:|||||||||||||||||||||||||||||
Db        273 EVKIGAKTSVIKEKDGKLVTGKDKGENDSSTDKGEGLVTAKEVIDAVNKAGWRMKTTTAN 332

Qy        122 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGW 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        333 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGW 392

Qy        182 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        393 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 452

Qy        242 SLGAGADAPTLSVDGD-ALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNR 300
              |||||||||||||| : |||||||  ||||||||||||||||||||||||||||||||||
Db        453 SLGAGADAPTLSVDDEGALNVGSKDANKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNH 512

Qy        301 IDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWII 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        513 IDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWII 572

Qy        361 KGTASGNSRGHFGASASVGYQW 382
              ||||||||||||||||||||||
Db        573 KGTASGNSRGHFGASASVGYQW 594


RESULT 8
US-09-669-974-9
; Sequence 9, Application US/09669974
; Patent No. 6333173
; GENERAL INFORMATION:
;  APPLICANT: PEAK, Ian Richard Anselm
;  APPLICANT:  JENNINGS, Michael Paul
;  APPLICANT:  MOXON, E. Richard
;  TITLE OF INVENTION: NOVEL SURFACE ANTIGEN
;  FILE REFERENCE: 065064/0128
;  CURRENT APPLICATION NUMBER: US/09/669,974
;  CURRENT FILING DATE:  2000-09-26
;  PRIOR APPLICATION NUMBER: US 09/377,155
;  PRIOR FILING DATE: 1999-08-19
;  PRIOR APPLICATION NUMBER: PCT/AU98/01031
;  PRIOR FILING DATE: 1998-12-14
;  PRIOR APPLICATION NUMBER: GB 9726398.2
;  PRIOR FILING DATE: 1997-12-12
;  NUMBER OF SEQ ID NOS: 33
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
;   LENGTH: 594
;   TYPE: PRT
;   ORGANISM: Neisseria meningitidis
US-09-669-974-9

  Query Match          96.9%;  Score 1886.5;  DB 4;  Length 594;
  Best Local Similarity   97.6%;  Pred. No. 5.1e-146;
  Matches  373;  Conservative    3;  Mismatches    5;  Indels    1;  Gaps    1;

Qy          2 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKT 61
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:|
Db        213 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKRT 272

Qy         62 EVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 121
              |||||||||||||||||||||||| |||||:|||||||||||||||||||||||||||||
Db        273 EVKIGAKTSVIKEKDGKLVTGKDKGENDSSTDKGEGLVTAKEVIDAVNKAGWRMKTTTAN 332

Qy        122 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGW 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        333 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGW 392

Qy        182 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        393 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 452

Qy        242 SLGAGADAPTLSVDGD-ALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNR 300
              |||||||||||||| : |||||||  ||||||||||||||||||||||||||||||||||
Db        453 SLGAGADAPTLSVDDEGALNVGSKDANKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNH 512

Qy        301 IDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWII 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        513 IDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWII 572

Qy        361 KGTASGNSRGHFGASASVGYQW 382
              ||||||||||||||||||||||
Db        573 KGTASGNSRGHFGASASVGYQW 594


RESULT 9
US-09-377-155-15
; Sequence 15, Application US/09377155
; Patent No. 6197312
; GENERAL INFORMATION:
;  APPLICANT:  PEAK, Ian Richard Anselm
;  APPLICANT:  JENNINGS, Michael Paul
;  APPLICANT:  MOXON, E. Richard
;  TITLE OF INVENTION:  NOVEL SURFACE ANTIGEN
;  FILE REFERENCE:  065064/0128
;  CURRENT APPLICATION NUMBER: US/09/377,155
;  CURRENT FILING DATE:  1999-08-19
;  PRIOR APPLICATION NUMBER:  PCT/AU98/01031
;  PRIOR FILING DATE:  1998-12-14
;  PRIOR APPLICATION NUMBER:  GB 9726398.2
;  PRIOR FILING DATE:  1997-12-12
;  NUMBER OF SEQ ID NOS:  33
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 15
;   LENGTH: 599
;   TYPE: PRT
;   ORGANISM: Neisseria meningitidis
US-09-377-155-15
```

```
  Query Match           96.6%;  Score 1881.5;  DB 3;  Length 599;
  Best Local Similarity  97.6%;  Pred. No. 1.3e-145;
  Matches  373;  Conservative    1;  Mismatches    7;  Indels    1;  Gaps    1;

Qy          2 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKT 61
              |||||||||||||||||||||||||||||| |||||||||||||||||||||||||:|
Db        218 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVHTYDTVEFLSADTKTTTVNVESKDNGKRT 277

Qy         62 EVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 121
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        278 EVKIGAKTSVIKEKDGKLVTGKGKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 337

Qy        122 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGW 181
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        338 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVKYDVNVGDALNVNQLQNSGW 397

Qy        182 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 241
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        398 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 457

Qy        242 SLGAGADAPTLSVDG-DALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNR 300
              ||||||||||||    |||||||  ||||||||||||||||||||||||||||||||||
Db        458 SLGAGADAPTLSVDDKGALNVGSKDANKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNR 517

Qy        301 IDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWII 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        518 IDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWII 577

Qy        361 KGTASGNSRGHFGASASVGYQW 382
              ||||||||||||||||||||||
Db        578 KGTASGNSRGHFGASASVGYQW 599
```

RESULT 10
US-09-669-974-15
; Sequence 15, Application US/09669974
; Patent No. 6333173
; GENERAL INFORMATION:
; APPLICANT: PEAK, Ian Richard Anselm
; APPLICANT: JENNINGS, Michael Paul
; APPLICANT: MOXON, E. Richard
; TITLE OF INVENTION: NOVEL SURFACE ANTIGEN
; FILE REFERENCE: 065064/0128
; CURRENT APPLICATION NUMBER: US/09/669,974
; CURRENT FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US 09/377,155
; PRIOR FILING DATE: 1999-08-19
; PRIOR APPLICATION NUMBER: PCT/AU98/01031
; PRIOR FILING DATE: 1998-12-14
; PRIOR APPLICATION NUMBER: GB 9726398.2
; PRIOR FILING DATE: 1997-12-12
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
; LENGTH: 599
; TYPE: PRT
; ORGANISM: Neisseria meningitidis
US-09-669-974-15

```
  Query Match           96.6%;  Score 1881.5;  DB 4;  Length 599;
  Best Local Similarity  97.6%;  Pred. No. 1.3e-145;
  Matches  373;  Conservative    1;  Mismatches    7;  Indels    1;  Gaps    1;

Qy          2 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKT 61
              |||||||||||||||||||||||||||||| |||||||||||||||||||||||||:|
Db        218 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVHTYDTVEFLSADTKTTTVNVESKDNGKRT 277

Qy         62 EVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 121
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        278 EVKIGAKTSVIKEKDGKLVTGKGKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 337

Qy        122 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGW 181
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        338 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVKYDVNVGDALNVNQLQNSGW 397

Qy        182 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 241
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        398 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 457

Qy        242 SLGAGADAPTLSVDG-DALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNR 300
              ||||||||||||    |||||||  ||||||||||||||||||||||||||||||||||
Db        458 SLGAGADAPTLSVDDKGALNVGSKDANKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNR 517

Qy        301 IDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWII 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        518 IDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWII 577

Qy        361 KGTASGNSRGHFGASASVGYQW 382
              ||||||||||||||||||||||
Db        578 KGTASGNSRGHFGASASVGYQW 599
```

RESULT 11
US-09-377-155-7
; Sequence 7, Application US/09377155
; Patent No. 6197312
; GENERAL INFORMATION:
; APPLICANT: PEAK, Ian Richard Anselm
; APPLICANT: JENNINGS, Michael Paul
; APPLICANT: MOXON, E. Richard
; TITLE OF INVENTION: NOVEL SURFACE ANTIGEN
; FILE REFERENCE: 065064/0128
; CURRENT APPLICATION NUMBER: US/09/377,155
; CURRENT FILING DATE: 1999-08-19
; PRIOR APPLICATION NUMBER: PCT/AU98/01031
; PRIOR FILING DATE: 1998-12-14
; PRIOR APPLICATION NUMBER: GB 9726398.2
; PRIOR FILING DATE: 1997-12-12
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 594
; TYPE: PRT
; ORGANISM: Neisseria meningitidis
US-09-377-155-7

```
  Query Match           96.3%;  Score 1874.5;  DB 3;  Length 594;
  Best Local Similarity  97.4%;  Pred. No. 4.9e-145;
  Matches  372;  Conservative    1;  Mismatches    8;  Indels    1;  Gaps    1;

Qy          2 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKT 61
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        213 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKT 272

Qy         62 EVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 121
              |||||||||||||||||||| | ||||||||||||||||||||||||||||||||||||
Db        273 EVKIGAKTSVIKEKDGKLVTGKGKDENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 332

Qy        122 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGW 181
              |||||||||||||||||||||||||||||||||||| |||||||||||||||||||||
Db        333 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVKYDVNVGDALNVNQLQNSGW 392

Qy        182 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 241
              |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        393 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMAPQFSSV 452

Qy        242 SLGAGADAPTLSVDGD-ALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNR 300
              |||||||||||||| : |||||||  ||||||||||||||||||||||||||||||||||
Db        453 SLGAGADAPTLSVDDEGALNVGSKDTNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNR 512

Qy        301 IDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWII 360
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        513 IDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGDTYRGEAGYAIGYSSISDGGNWII 572

Qy        361 KGTASGNSRGHFGASASVGYQW 382
              ||||||||||||||||||||||
```

```
Db        573 KGTASGNSRGHFGASASVGYQW 594


RESULT 12
US-09-669-974-7
; Sequence 7, Application US/09669974
; Patent No. 6333173
; GENERAL INFORMATION:
;  APPLICANT: PEAK, Ian Richard Anselm
;  APPLICANT:  JENNINGS, Michael Paul
;  APPLICANT:  MOXON, E. Richard
;  TITLE OF INVENTION: NOVEL SURFACE ANTIGEN
;  FILE REFERENCE: 065064/0128
;  CURRENT APPLICATION NUMBER: US/09/669,974
;  CURRENT FILING DATE:  2000-09-26
;  PRIOR APPLICATION NUMBER: US 09/377,155
;  PRIOR FILING DATE: 1999-08-19
;  PRIOR APPLICATION NUMBER: PCT/AU98/01031
;  PRIOR FILING DATE: 1998-12-14
;  PRIOR APPLICATION NUMBER: GB 9726398.2
;  PRIOR FILING DATE: 1997-12-12
;  NUMBER OF SEQ ID NOS: 33
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
;   LENGTH: 594
;   TYPE: PRT
;   ORGANISM: Neisseria meningitidis
US-09-669-974-7

  Query Match             96.3%;  Score 1874.5;  DB 4;  Length 594;
  Best Local Similarity   97.4%;  Pred. No. 4.9e-145;
  Matches  372;  Conservative    1;  Mismatches    8;  Indels    1;  Gaps    1;

Qy          2 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKT 61
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        213 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKT 272

Qy         62 EVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 121
              ||||||||||||||||||||||| | ||||||||||||||||||||||||||||||||||
Db        273 EVKIGAKTSVIKEKDGKLVTGKGKDENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 332

Qy        122 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGW 181
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        333 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVKYDVNVGDALNVNQLQNSGW 392

Qy        182 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 241
              |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        393 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMAPQFSSV 452

Qy        242 SLGAGADAPTLSVDGD-ALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNR 300
              ||||||||||||||  : |||||||  ||||||||||||||||||||||||||||||||||
Db        453 SLGAGADAPTLSVDDEGALNVGSKDTNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNR 512

Qy        301 IDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWII 360
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        513 IDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGDTYRGEAGYAIGYSSISDGGNWII 572

Qy        361 KGTASGNSRGHFGASASVGYQW 382
              ||||||||||||||||||||||
Db        573 KGTASGNSRGHFGASASVGYQW 594


RESULT 13
US-09-377-155-17
; Sequence 17, Application US/09377155
; Patent No. 6197312
; GENERAL INFORMATION:
;  APPLICANT:  PEAK, Ian Richard Anselm
;  APPLICANT:  JENNINGS, Michael Paul
;  APPLICANT:  MOXON, E. Richard
;  TITLE OF INVENTION:  NOVEL SURFACE ANTIGEN
;  FILE REFERENCE:  065064/0128
;  CURRENT APPLICATION NUMBER: US/09/377,155
;  CURRENT FILING DATE:  1999-08-19
;  PRIOR APPLICATION NUMBER:  PCT/AU98/01031
;  PRIOR FILING DATE:  1998-12-14
;  PRIOR APPLICATION NUMBER:  GB 9726398.2
;  PRIOR FILING DATE:  1997-12-12
;  NUMBER OF SEQ ID NOS:  33
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 17
;   LENGTH: 592
;   TYPE: PRT
;   ORGANISM: Neisseria meningitidis
US-09-377-155-17

  Query Match             95.6%;  Score 1861.5;  DB 3;  Length 592;
  Best Local Similarity   96.9%;  Pred. No. 5.6e-144;
  Matches  370;  Conservative    2;  Mismatches    9;  Indels    1;  Gaps    1;

Qy          2 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKT 61
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        211 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKT 270

Qy         62 EVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 121
              ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        271 EVKIGAKTSVIKEKDGKLVTGKGKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 330

Qy        122 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGW 181
              |||||||||||||||| |||||| |||||||||||||||| ||||||||||||||||||||
Db        331 GQTGQADKFETVTSGTKVTFASGNGTTATVSKDDQGNITVKYDVNVGDALNVNQLQNSGW 390

Qy        182 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        391 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 450

Qy        242 SLGAGADAPTLSVDGD-ALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNR 300
              ||||||||||||||  : |||||||  ||||||||||||||||||||||||||||||||||
Db        451 SLGAGADAPTLSVDDEGALNVGSKDANKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNR 510

Qy        301 IDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWII 360
              ||||:|||||||||||||||||||||||||||||| |||||||||||| |||||||||||
Db        511 IDNVNGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYLGEAGYAIGYSSISAGGNWII 570

Qy        361 KGTASGNSRGHFGASASVGYQW 382
              ||||||||||||||||||||||
Db        571 KGTASGNSRGHFGASASVGYQW 592


RESULT 14
US-09-669-974-17
; Sequence 17, Application US/09669974
; Patent No. 6333173
; GENERAL INFORMATION:
;  APPLICANT: PEAK, Ian Richard Anselm
;  APPLICANT:  JENNINGS, Michael Paul
;  APPLICANT:  MOXON, E. Richard
;  TITLE OF INVENTION: NOVEL SURFACE ANTIGEN
;  FILE REFERENCE: 065064/0128
;  CURRENT APPLICATION NUMBER: US/09/669,974
;  CURRENT FILING DATE:  2000-09-26
;  PRIOR APPLICATION NUMBER: US 09/377,155
;  PRIOR FILING DATE: 1999-08-19
;  PRIOR APPLICATION NUMBER: PCT/AU98/01031
;  PRIOR FILING DATE: 1998-12-14
;  PRIOR APPLICATION NUMBER: GB 9726398.2
;  PRIOR FILING DATE: 1997-12-12
;  NUMBER OF SEQ ID NOS: 33
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17
;   LENGTH: 592
;   TYPE: PRT
;   ORGANISM: Neisseria meningitidis
US-09-669-974-17

  Query Match             95.6%;  Score 1861.5;  DB 4;  Length 592;
```

```
  Best Local Similarity   96.9%;  Pred. No. 5.6e-144;
  Matches  370;  Conservative    2;  Mismatches    9;  Indels    1;  Gaps    1;

Qy          2 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKT 61
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        211 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKT 270

Qy         62 EVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 121
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        271 EVKIGAKTSVIKEKDGKLVTGKGKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 330

Qy        122 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGW 181
              |||||||||||||||| |||||| |||||||||||||||| ||||||||||||||||||||
Db        331 GQTGQADKFETVTSGTKVTFASGNGTTATVSKDDQGNITVKYDVNVGDALNVNQLQNSGW 390

Qy        182 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        391 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 450

Qy        242 SLGAGADAPTLSVDGD-ALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNR 300
              |||||||||||||| : |||||||  ||||||||||||||||||||||||||||||||||||
Db        451 SLGAGADAPTLSVDDEGALNVGSKDANKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNR 510

Qy        301 IDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWII 360
              ||||:|||||||||||||||||||||||||||||||| |||||||||||| ||||||||||
Db        511 IDNVNGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYLGEAGYAIGYSSISAGGNWII 570

Qy        361 KGTASGNSRGHFGASASVGYQW 382
              ||||||||||||||||||||||
Db        571 KGTASGNSRGHFGASASVGYQW 592


RESULT 15
US-09-377-155-5
; Sequence 5, Application US/09377155
; Patent No. 6197312
; GENERAL INFORMATION:
;  APPLICANT:  PEAK, Ian Richard Anselm
;  APPLICANT:  JENNINGS, Michael Paul
;  APPLICANT:  MOXON, E. Richard
;  TITLE OF INVENTION:  NOVEL SURFACE ANTIGEN
;  FILE REFERENCE:  065064/0128
;  CURRENT APPLICATION NUMBER: US/09/377,155
;  CURRENT FILING DATE:  1999-08-19
;  PRIOR APPLICATION NUMBER:  PCT/AU98/01031
;  PRIOR FILING DATE:  1998-12-14
;  PRIOR APPLICATION NUMBER:  GB 9726398.2
;  PRIOR FILING DATE:  1997-12-12
;  NUMBER OF SEQ ID NOS:  33
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 5
;   LENGTH: 598
;   TYPE: PRT
;   ORGANISM: Neisseria meningitidis
US-09-377-155-5

  Query Match           95.6%;  Score 1861.5;  DB 3;  Length 598;
  Best Local Similarity   96.3%;  Pred. No. 5.7e-144;
  Matches  368;  Conservative    3;  Mismatches   10;  Indels    1;  Gaps    1;

Qy          2 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKT 61
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:|
Db        217 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKRT 276

Qy         62 EVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 121
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        277 EVKIGAKTSVIKEKDGKLVTGKGKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 336

Qy        122 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGW 181
              |||||||||||||||| |||||| |||||||||||||||| ||||||||||||||||||||
Db        337 GQTGQADKFETVTSGTKVTFASGNGTTATVSKDDQGNITVKYDVNVGDALNVNQLQNSGW 396

Qy        182 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        397 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 456

Qy        242 SLGAGADAPTLSVDGD-ALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNR 300
              |||||||||||||| : |||||||  ||||||||||||||||||||||||||||||||||||
Db        457 SLGAGADAPTLSVDDEGALNVGSKDANKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNR 516

Qy        301 IDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWII 360
              |||||||||||||||||||| |||||||||||||||||||||||||||||||| |||:|
Db        517 IDNVDGNARAGIAQAIATAGLAQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDTGNWVI 576

Qy        361 KGTASGNSRGHFGASASVGYQW 382
              |||||||||||| ||||||||
Db        577 KGTASGNSRGHFGTSASVGYQW 598


Search completed: October  6, 2003, 09:36:00
Job time : 11.2154 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: October 6, 2003, 09:31:05 ; Search time 19.136 Seconds
(without alignments)
3158.312 Million cell updates/sec

Title: US-09-771-382-38
Perfect score: 1947
Sequence: 1 NRAASVKDVLNAGWNIKGVK.....TASGNSRGHFGASASVGYQW 382

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 587654 seqs, 158212981 residues

Total number of hits satisfying chosen parameters: 587654

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1947	100.0	382	10	US-09-771-382-38	Sequence 38, Appl
2	1947	100.0	433	10	US-09-771-382-26	Sequence 26, Appl
3	1941	99.7	461	10	US-09-771-382-35	Sequence 35, Appl
4	1941	99.7	512	10	US-09-771-382-23	Sequence 23, Appl
5	1941	99.7	540	10	US-09-771-382-33	Sequence 33, Appl
6	1941	99.7	591	10	US-09-797-862-11	Sequence 11, Appl
7	1941	99.7	591	10	US-09-797-862-21	Sequence 21, Appl
8	1941	99.7	591	10	US-09-771-382-1	Sequence 1, Appli
9	1941	99.7	591	10	US-09-771-382-5	Sequence 5, Appli
10	1937	99.5	592	10	US-09-797-862-2	Sequence 2, Appli
11	1891.5	97.1	502	10	US-09-771-382-27	Sequence 27, Appl
12	1886.5	96.9	594	10	US-09-797-862-9	Sequence 9, Appli
13	1886.5	96.9	594	10	US-09-771-382-4	Sequence 4, Appli
14	1881.5	96.6	599	10	US-09-797-862-15	Sequence 15, Appl
15	1881.5	96.6	599	10	US-09-771-382-6	Sequence 6, Appli
16	1874.5	96.3	594	10	US-09-797-862-7	Sequence 7, Appli
17	1874.5	96.3	594	10	US-09-771-382-9	Sequence 9, Appli
18	1861.5	95.6	462	10	US-09-771-382-36	Sequence 36, Appl
19	1861.5	95.6	513	10	US-09-771-382-24	Sequence 24, Appl
20	1861.5	95.6	541	10	US-09-771-382-34	Sequence 34, Appl
21	1861.5	95.6	592	10	US-09-797-862-17	Sequence 17, Appl
22	1861.5	95.6	592	10	US-09-771-382-2	Sequence 2, Appli
23	1861.5	95.6	598	10	US-09-797-862-5	Sequence 5, Appli
24	1861.5	95.6	598	10	US-09-797-862-13	Sequence 13, Appl
25	1861.5	95.6	598	10	US-09-771-382-7	Sequence 7, Appli
26	1861.5	95.6	598	10	US-09-771-382-8	Sequence 8, Appli
27	1854.5	95.2	592	10	US-09-771-382-10	Sequence 10, Appl
28	1817.5	93.3	589	10	US-09-797-862-19	Sequence 19, Appl
29	1817.5	93.3	589	10	US-09-771-382-3	Sequence 3, Appli
30	1810	93.0	407	10	US-09-771-382-25	Sequence 25, Appl
31	1808	92.9	356	10	US-09-771-382-37	Sequence 37, Appl
32	1736.5	89.2	604	10	US-09-771-382-11	Sequence 11, Appl
33	710	36.5	1098	10	US-09-797-862-32	Sequence 32, Appl
34	660	33.9	2353	10	US-09-797-862-33	Sequence 33, Appl
35	607.5	31.2	201	10	US-09-771-382-39	Sequence 39, Appl
36	318.5	16.4	1778	12	US-10-238-075-749	Sequence 749, App
37	298	15.3	2122	10	US-09-813-214A-9	Sequence 9, Appli
38	297	15.3	1833	12	US-10-175-275-4	Sequence 4, Appli
39	297	15.3	1833	12	US-10-175-282-4	Sequence 4, Appli
40	297	15.3	1992	12	US-10-175-275-3	Sequence 3, Appli
41	297	15.3	1992	12	US-10-175-282-3	Sequence 3, Appli
42	234.5	12.0	2039	15	US-10-192-584-7	Sequence 7, Appli
43	227.5	11.7	2042	15	US-10-192-584-6	Sequence 6, Appli
44	186.5	9.6	915	12	US-10-193-764-35	Sequence 35, Appl
45	186.5	9.6	1222	12	US-10-193-764-37	Sequence 37, Appl

ALIGNMENTS

RESULT 1
US-09-771-382-38
; Sequence 38, Application US/09771382
; Patent No. US20020160016A1
; GENERAL INFORMATION:
; APPLICANT: Peak, Ian
; APPLICANT: Jennings, Michael
; TITLE OF INVENTION: MODIFIED SURFACE ANTIGEN
; FILE REFERENCE: 8795-24U1
; CURRENT APPLICATION NUMBER: US/09/771,382
; CURRENT FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: US 60/177,917
; PRIOR FILING DATE: 2000-01-25
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 38
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Neisseria meningitidis
US-09-771-382-38

Query Match 100.0%; Score 1947; DB 10; Length 382;
Best Local Similarity 100.0%; Pred. No. 1.1e-149;
Matches 382; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 NRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKK 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 NRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKK 60

Qy         61 TEVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TEVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTA 120

Qy        121 NGQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 NGQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSG 180

Qy        181 WNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSS 240
```

```
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 WNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSS 240

Qy        241 VSLGAGADAPTLSVDGDALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNR 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 VSLGAGADAPTLSVDGDALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNR 300

Qy        301 IDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWII 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 IDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWII 360

Qy        361 KGTASGNSRGHFGASASVGYQW 382
              ||||||||||||||||||||||
Db        361 KGTASGNSRGHFGASASVGYQW 382


RESULT 2
US-09-771-382-26
; Sequence 26, Application US/09771382
; Patent No. US20020160016A1
; GENERAL INFORMATION:
;  APPLICANT: Peak, Ian
;  APPLICANT:  Jennings, Michael
;  TITLE OF INVENTION: MODIFIED SURFACE ANTIGEN
;  FILE REFERENCE: 8795-24U1
;  CURRENT APPLICATION NUMBER: US/09/771,382
;  CURRENT FILING DATE:  2001-01-25
;  PRIOR APPLICATION NUMBER: US 60/177,917
;  PRIOR FILING DATE: 2000-01-25
;  NUMBER OF SEQ ID NOS: 52
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 26
;   LENGTH: 433
;   TYPE: PRT
;   ORGANISM: Neisseria meningitidis
US-09-771-382-26

  Query Match             100.0%;  Score 1947;  DB 10;  Length 433;
  Best Local Similarity   100.0%;  Pred. No. 1.3e-149;
  Matches  382;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 NRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKK 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         52 NRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKK 111

Qy         61 TEVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        112 TEVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTA 171

Qy        121 NGQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        172 NGQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSG 231

Qy        181 WNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSS 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        232 WNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSS 291

Qy        241 VSLGAGADAPTLSVDGDALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNR 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        292 VSLGAGADAPTLSVDGDALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNR 351

Qy        301 IDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWII 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        352 IDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWII 411

Qy        361 KGTASGNSRGHFGASASVGYQW 382
              ||||||||||||||||||||||
Db        412 KGTASGNSRGHFGASASVGYQW 433


RESULT 3
US-09-771-382-35
; Sequence 35, Application US/09771382
; Patent No. US20020160016A1
; GENERAL INFORMATION:
;  APPLICANT: Peak, Ian
;  APPLICANT:  Jennings, Michael
;  TITLE OF INVENTION: MODIFIED SURFACE ANTIGEN
;  FILE REFERENCE: 8795-24U1
;  CURRENT APPLICATION NUMBER: US/09/771,382
;  CURRENT FILING DATE:  2001-01-25
;  PRIOR APPLICATION NUMBER: US 60/177,917
;  PRIOR FILING DATE: 2000-01-25
;  NUMBER OF SEQ ID NOS: 52
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 35
;   LENGTH: 461
;   TYPE: PRT
;   ORGANISM: Neisseria meningitidis
US-09-771-382-35

  Query Match             99.7%;  Score 1941;  DB 10;  Length 461;
  Best Local Similarity   100.0%;  Pred. No. 4.3e-149;
  Matches  381;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKT 61
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKT 140

Qy         62 EVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 EVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 200

Qy        122 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGW 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGW 260

Qy        182 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 320

Qy        242 SLGAGADAPTLSVDGDALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRI 301
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 SLGAGADAPTLSVDGDALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRI 380

Qy        302 DNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIK 361
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        381 DNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIK 440

Qy        362 GTASGNSRGHFGASASVGYQW 382
              |||||||||||||||||||||
Db        441 GTASGNSRGHFGASASVGYQW 461


RESULT 4
US-09-771-382-23
; Sequence 23, Application US/09771382
; Patent No. US20020160016A1
; GENERAL INFORMATION:
;  APPLICANT: Peak, Ian
;  APPLICANT:  Jennings, Michael
;  TITLE OF INVENTION: MODIFIED SURFACE ANTIGEN
;  FILE REFERENCE: 8795-24U1
;  CURRENT APPLICATION NUMBER: US/09/771,382
;  CURRENT FILING DATE:  2001-01-25
;  PRIOR APPLICATION NUMBER: US 60/177,917
;  PRIOR FILING DATE: 2000-01-25
;  NUMBER OF SEQ ID NOS: 52
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 23
;   LENGTH: 512
;   TYPE: PRT
;   ORGANISM: Neisseria meningitidis
US-09-771-382-23

  Query Match             99.7%;  Score 1941;  DB 10;  Length 512;
```

```
  Best Local Similarity   100.0%;  Pred. No. 5e-149;
  Matches  381;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKT 61
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        132 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKT 191

Qy         62 EVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        192 EVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 251

Qy        122 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGW 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        252 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGW 311

Qy        182 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        312 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 371

Qy        242 SLGAGADAPTLSVDGDALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRI 301
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        372 SLGAGADAPTLSVDGDALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRI 431

Qy        302 DNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIK 361
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        432 DNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIK 491

Qy        362 GTASGNSRGHFGASASVGYQW 382
              |||||||||||||||||||||
Db        492 GTASGNSRGHFGASASVGYQW 512


RESULT 5
US-09-771-382-33
; Sequence 33, Application US/09771382
; Patent No. US20020160016A1
; GENERAL INFORMATION:
;  APPLICANT: Peak, Ian
;  APPLICANT:  Jennings, Michael
;  TITLE OF INVENTION: MODIFIED SURFACE ANTIGEN
;  FILE REFERENCE: 8795-24U1
;  CURRENT APPLICATION NUMBER: US/09/771,382
;  CURRENT FILING DATE:  2001-01-25
;  PRIOR APPLICATION NUMBER: US 60/177,917
;  PRIOR FILING DATE: 2000-01-25
;  NUMBER OF SEQ ID NOS: 52
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 33
;   LENGTH: 540
;   TYPE: PRT
;   ORGANISM: Neisseria meningitidis
US-09-771-382-33

  Query Match             99.7%;  Score 1941;  DB 10;  Length 540;
  Best Local Similarity   100.0%;  Pred. No. 5.4e-149;
  Matches  381;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKT 61
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        160 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKT 219

Qy         62 EVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        220 EVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 279

Qy        122 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGW 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        280 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGW 339

Qy        182 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        340 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 399

Qy        242 SLGAGADAPTLSVDGDALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRI 301
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        400 SLGAGADAPTLSVDGDALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRI 459

Qy        302 DNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIK 361
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        460 DNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIK 519

Qy        362 GTASGNSRGHFGASASVGYQW 382
              |||||||||||||||||||||
Db        520 GTASGNSRGHFGASASVGYQW 540


RESULT 6
US-09-797-862-11
; Sequence 11, Application US/09797862
; Patent No. US20020102276A1
; GENERAL INFORMATION:
;  APPLICANT: PEAK, IAN RICHARD ANSELM
;  APPLICANT:  JENNINGS, MICHAEL PAUL
;  APPLICANT:  MOXON, E. RICHARD
;  TITLE OF INVENTION: NOVEL SURFACE ANTIGEN
;  FILE REFERENCE: 065064/0134
;  CURRENT APPLICATION NUMBER: US/09/797,862
;  CURRENT FILING DATE:  2001-05-03
;  PRIOR APPLICATION NUMBER: PCT/AU98/01031
;  PRIOR FILING DATE: 1998-12-14
;  PRIOR APPLICATION NUMBER: GB 9726398.2
;  PRIOR FILING DATE: 1997-12-12
;  NUMBER OF SEQ ID NOS: 33
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
;   LENGTH: 591
;   TYPE: PRT
;   ORGANISM: Neisseria meningitidis
US-09-797-862-11

  Query Match             99.7%;  Score 1941;  DB 10;  Length 591;
  Best Local Similarity   100.0%;  Pred. No. 6.1e-149;
  Matches  381;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKT 61
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        211 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKT 270

Qy         62 EVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        271 EVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 330

Qy        122 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGW 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        331 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGW 390

Qy        182 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        391 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 450

Qy        242 SLGAGADAPTLSVDGDALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRI 301
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        451 SLGAGADAPTLSVDGDALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRI 510

Qy        302 DNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIK 361
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        511 DNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIK 570

Qy        362 GTASGNSRGHFGASASVGYQW 382
              |||||||||||||||||||||
Db        571 GTASGNSRGHFGASASVGYQW 591


RESULT 7
US-09-797-862-21
; Sequence 21, Application US/09797862
```

```
; Patent No. US20020102276A1
; GENERAL INFORMATION:
;  APPLICANT: PEAK, IAN RICHARD ANSELM
;  APPLICANT:  JENNINGS, MICHAEL PAUL
;  APPLICANT:  MOXON, E. RICHARD
;  TITLE OF INVENTION: NOVEL SURFACE ANTIGEN
;  FILE REFERENCE: 065064/0134
;  CURRENT APPLICATION NUMBER: US/09/797,862
;  CURRENT FILING DATE:  2001-05-03
;  PRIOR APPLICATION NUMBER: PCT/AU98/01031
;  PRIOR FILING DATE: 1998-12-14
;  PRIOR APPLICATION NUMBER: GB 9726398.2
;  PRIOR FILING DATE: 1997-12-12
;  NUMBER OF SEQ ID NOS: 33
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 21
;   LENGTH: 591
;   TYPE: PRT
;   ORGANISM: Neisseria meningitidis
US-09-797-862-21

  Query Match            99.7%;  Score 1941;  DB 10;  Length 591;
  Best Local Similarity  100.0%;  Pred. No. 6.1e-149;
  Matches  381;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKT 61
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        211 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKT 270

Qy         62 EVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        271 EVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 330

Qy        122 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGW 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        331 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGW 390

Qy        182 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        391 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 450

Qy        242 SLGAGADAPTLSVDGDALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRI 301
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        451 SLGAGADAPTLSVDGDALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRI 510

Qy        302 DNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIK 361
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        511 DNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIK 570

Qy        362 GTASGNSRGHFGASASVGYQW 382
              |||||||||||||||||||||
Db        571 GTASGNSRGHFGASASVGYQW 591


RESULT 8
US-09-771-382-1
; Sequence 1, Application US/09771382
; Patent No. US20020160016A1
; GENERAL INFORMATION:
;  APPLICANT: Peak, Ian
;  APPLICANT:  Jennings, Michael
;  TITLE OF INVENTION: MODIFIED SURFACE ANTIGEN
;  FILE REFERENCE: 8795-24U1
;  CURRENT APPLICATION NUMBER: US/09/771,382
;  CURRENT FILING DATE:  2001-01-25
;  PRIOR APPLICATION NUMBER: US 60/177,917
;  PRIOR FILING DATE: 2000-01-25
;  NUMBER OF SEQ ID NOS: 52
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
;   LENGTH: 591
;   TYPE: PRT
;   ORGANISM: Neisseria meningitidis
US-09-771-382-1

  Query Match            99.7%;  Score 1941;  DB 10;  Length 591;
  Best Local Similarity  100.0%;  Pred. No. 6.1e-149;
  Matches  381;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKT 61
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        211 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKT 270

Qy         62 EVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        271 EVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 330

Qy        122 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGW 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        331 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGW 390

Qy        182 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        391 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 450

Qy        242 SLGAGADAPTLSVDGDALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRI 301
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        451 SLGAGADAPTLSVDGDALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRI 510

Qy        302 DNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIK 361
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        511 DNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIK 570

Qy        362 GTASGNSRGHFGASASVGYQW 382
              |||||||||||||||||||||
Db        571 GTASGNSRGHFGASASVGYQW 591


RESULT 9
US-09-771-382-5
; Sequence 5, Application US/09771382
; Patent No. US20020160016A1
; GENERAL INFORMATION:
;  APPLICANT: Peak, Ian
;  APPLICANT:  Jennings, Michael
;  TITLE OF INVENTION: MODIFIED SURFACE ANTIGEN
;  FILE REFERENCE: 8795-24U1
;  CURRENT APPLICATION NUMBER: US/09/771,382
;  CURRENT FILING DATE:  2001-01-25
;  PRIOR APPLICATION NUMBER: US 60/177,917
;  PRIOR FILING DATE: 2000-01-25
;  NUMBER OF SEQ ID NOS: 52
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
;   LENGTH: 591
;   TYPE: PRT
;   ORGANISM: Neisseria meningitidis
US-09-771-382-5

  Query Match            99.7%;  Score 1941;  DB 10;  Length 591;
  Best Local Similarity  100.0%;  Pred. No. 6.1e-149;
  Matches  381;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKT 61
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        211 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKT 270

Qy         62 EVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        271 EVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 330

Qy        122 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGW 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        331 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGW 390

Qy        182 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 241
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        391 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 450

Qy        242 SLGAGADAPTLSVDGDALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRI 301
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        451 SLGAGADAPTLSVDGDALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRI 510

Qy        302 DNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIK 361
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        511 DNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIK 570

Qy        362 GTASGNSRGHFGASASVGYQW 382
                |||||||||||||||||||||
Db        571 GTASGNSRGHFGASASVGYQW 591


RESULT 10
US-09-797-862-2
; Sequence 2, Application US/09797862
; Patent No. US20020102276A1
; GENERAL INFORMATION:
;  APPLICANT: PEAK, IAN RICHARD ANSELM
;  APPLICANT:  JENNINGS, MICHAEL PAUL
;  APPLICANT:  MOXON, E. RICHARD
;  TITLE OF INVENTION: NOVEL SURFACE ANTIGEN
;  FILE REFERENCE: 065064/0134
;  CURRENT APPLICATION NUMBER: US/09/797,862
;  CURRENT FILING DATE:  2001-05-03
;  PRIOR APPLICATION NUMBER: PCT/AU98/01031
;  PRIOR FILING DATE: 1998-12-14
;  PRIOR APPLICATION NUMBER: GB 9726398.2
;  PRIOR FILING DATE: 1997-12-12
;  NUMBER OF SEQ ID NOS: 33
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 592
;   TYPE: PRT
;   ORGANISM: Neisseria meningitidis
US-09-797-862-2

  Query Match             99.5%;  Score 1937;  DB 10;  Length 592;
  Best Local Similarity   99.7%;  Pred. No. 1.3e-148;
  Matches  380;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKT 61
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        212 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKT 271

Qy         62 EVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 121
                ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        272 EVKIGVKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 331

Qy        122 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGW 181
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        332 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGW 391

Qy        182 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 241
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        392 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 451

Qy        242 SLGAGADAPTLSVDGDALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRI 301
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        452 SLGAGADAPTLSVDGDALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRI 511

Qy        302 DNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIK 361
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        512 DNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIK 571

Qy        362 GTASGNSRGHFGASASVGYQW 382
                |||||||||||||||||||||
Db        572 GTASGNSRGHFGASASVGYQW 592


RESULT 11
US-09-771-382-27
; Sequence 27, Application US/09771382
; Patent No. US20020160016A1
; GENERAL INFORMATION:
;  APPLICANT: Peak, Ian
;  APPLICANT:  Jennings, Michael
;  TITLE OF INVENTION: MODIFIED SURFACE ANTIGEN
;  FILE REFERENCE: 8795-24U1
;  CURRENT APPLICATION NUMBER: US/09/771,382
;  CURRENT FILING DATE:  2001-01-25
;  PRIOR APPLICATION NUMBER: US 60/177,917
;  PRIOR FILING DATE: 2000-01-25
;  NUMBER OF SEQ ID NOS: 52
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 27
;   LENGTH: 502
;   TYPE: PRT
;   ORGANISM: Neisseria meningitidis
US-09-771-382-27

  Query Match             97.1%;  Score 1891.5;  DB 10;  Length 502;
  Best Local Similarity   97.9%;  Pred. No. 5e-145;
  Matches  374;  Conservative    1;  Mismatches    0;  Indels    7;  Gaps    1;

Qy          1 NRAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKK 60
                :|||||||||||||||||||       ||||||||||||||||||||||||||||||||
Db        128 DRAASVKDVLNAGWNIKGVK-------NVDFVRTYDTVEFLSADTKTTTVNVESKDNGKK 180

Qy         61 TEVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTA 120
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TEVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTA 240

Qy        121 NGQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSG 180
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 NGQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSG 300

Qy        181 WNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSS 240
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 WNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSS 360

Qy        241 VSLGAGADAPTLSVDGDALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNR 300
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 VSLGAGADAPTLSVDGDALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNR 420

Qy        301 IDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWII 360
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 IDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWII 480

Qy        361 KGTASGNSRGHFGASASVGYQW 382
                |||||||||||||||||||||
Db        481 KGTASGNSRGHFGASASVGYQW 502


RESULT 12
US-09-797-862-9
; Sequence 9, Application US/09797862
; Patent No. US20020102276A1
; GENERAL INFORMATION:
;  APPLICANT: PEAK, IAN RICHARD ANSELM
;  APPLICANT:  JENNINGS, MICHAEL PAUL
;  APPLICANT:  MOXON, E. RICHARD
;  TITLE OF INVENTION: NOVEL SURFACE ANTIGEN
;  FILE REFERENCE: 065064/0134
;  CURRENT APPLICATION NUMBER: US/09/797,862
;  CURRENT FILING DATE:  2001-05-03
;  PRIOR APPLICATION NUMBER: PCT/AU98/01031
;  PRIOR FILING DATE: 1998-12-14
;  PRIOR APPLICATION NUMBER: GB 9726398.2
;  PRIOR FILING DATE: 1997-12-12
;  NUMBER OF SEQ ID NOS: 33
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
```

```
;    LENGTH: 594
;    TYPE: PRT
;    ORGANISM: Neisseria meningitidis
US-09-797-862-9

  Query Match             96.9%;  Score 1886.5;  DB 10;  Length 594;
  Best Local Similarity   97.6%;  Pred. No. 1.6e-144;
  Matches  373;  Conservative    3;  Mismatches    5;  Indels    1;  Gaps    1;

Qy          2 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKT 61
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:|
Db        213 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKRT 272

Qy         62 EVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 121
              ||||||||||||||||||||||||||| ||||:|||||||||||||||||||||||||||
Db        273 EVKIGAKTSVIKEKDGKLVTGKDKGENDSSTDKGEGLVTAKEVIDAVNKAGWRMKTTTAN 332

Qy        122 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGW 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        333 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGW 392

Qy        182 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        393 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 452

Qy        242 SLGAGADAPTLSVDGD-ALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNR 300
              |||||||||||||| : |||||||  ||||||||||||||||||||||||||||||||| 
Db        453 SLGAGADAPTLSVDDEGALNVGSKDANKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNH 512

Qy        301 IDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWII 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        513 IDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWII 572

Qy        361 KGTASGNSRGHFGASASVGYQW 382
              ||||||||||||||||||||||
Db        573 KGTASGNSRGHFGASASVGYQW 594


RESULT 13
US-09-771-382-4
; Sequence 4, Application US/09771382
; Patent No. US20020160016A1
; GENERAL INFORMATION:
;  APPLICANT: Peak, Ian
;  APPLICANT:  Jennings, Michael
;  TITLE OF INVENTION: MODIFIED SURFACE ANTIGEN
;  FILE REFERENCE: 8795-24U1
;  CURRENT APPLICATION NUMBER: US/09/771,382
;  CURRENT FILING DATE:  2001-01-25
;  PRIOR APPLICATION NUMBER: US 60/177,917
;  PRIOR FILING DATE: 2000-01-25
;  NUMBER OF SEQ ID NOS: 52
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
;   LENGTH: 594
;   TYPE: PRT
;   ORGANISM: Neisseria meningitidis
US-09-771-382-4

  Query Match             96.9%;  Score 1886.5;  DB 10;  Length 594;
  Best Local Similarity   97.6%;  Pred. No. 1.6e-144;
  Matches  373;  Conservative    3;  Mismatches    5;  Indels    1;  Gaps    1;

Qy          2 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKT 61
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:|
Db        213 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKRT 272

Qy         62 EVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 121
              ||||||||||||||||||||||||||| ||||:|||||||||||||||||||||||||||
Db        273 EVKIGAKTSVIKEKDGKLVTGKDKGENDSSTDKGEGLVTAKEVIDAVNKAGWRMKTTTAN 332

Qy        122 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGW 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        333 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGW 392

Qy        182 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        393 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 452

Qy        242 SLGAGADAPTLSVDGD-ALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNR 300
              |||||||||||||| : |||||||  ||||||||||||||||||||||||||||||||| 
Db        453 SLGAGADAPTLSVDDEGALNVGSKDANKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNH 512

Qy        301 IDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWII 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        513 IDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWII 572

Qy        361 KGTASGNSRGHFGASASVGYQW 382
              ||||||||||||||||||||||
Db        573 KGTASGNSRGHFGASASVGYQW 594


RESULT 14
US-09-797-862-15
; Sequence 15, Application US/09797862
; Patent No. US20020102276A1
; GENERAL INFORMATION:
;  APPLICANT: PEAK, IAN RICHARD ANSELM
;  APPLICANT:  JENNINGS, MICHAEL PAUL
;  APPLICANT:  MOXON, E. RICHARD
;  TITLE OF INVENTION: NOVEL SURFACE ANTIGEN
;  FILE REFERENCE: 065064/0134
;  CURRENT APPLICATION NUMBER: US/09/797,862
;  CURRENT FILING DATE:  2001-05-03
;  PRIOR APPLICATION NUMBER: PCT/AU98/01031
;  PRIOR FILING DATE: 1998-12-14
;  PRIOR APPLICATION NUMBER: GB 9726398.2
;  PRIOR FILING DATE: 1997-12-12
;  NUMBER OF SEQ ID NOS: 33
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
;   LENGTH: 599
;   TYPE: PRT
;   ORGANISM: Neisseria meningitidis
US-09-797-862-15

  Query Match             96.6%;  Score 1881.5;  DB 10;  Length 599;
  Best Local Similarity   97.6%;  Pred. No. 4.1e-144;
  Matches  373;  Conservative    1;  Mismatches    7;  Indels    1;  Gaps    1;

Qy          2 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKT 61
              ||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||:|
Db        218 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVHTYDTVEFLSADTKTTTVNVESKDNGKRT 277

Qy         62 EVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 121
              |||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||
Db        278 EVKIGAKTSVIKEKDGKLVTGKGKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 337

Qy        122 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGW 181
              |||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db        338 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVKYDVNVGDALNVNQLQNSGW 397

Qy        182 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        398 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 457

Qy        242 SLGAGADAPTLSVDG-DALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNR 300
              ||||||||||||||   |||||||  ||||||||||||||||||||||||||||||||||
Db        458 SLGAGADAPTLSVDDKGALNVGSKDANKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNR 517

Qy        301 IDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWII 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        518 IDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWII 577

Qy        361 KGTASGNSRGHFGASASVGYQW 382
              ||||||||||||||||||||||
```

```
Db        578 KGTASGNSRGHFGASASVGYQW 599


RESULT 15
US-09-771-382-6
; Sequence 6, Application US/09771382
; Patent No. US20020160016A1
; GENERAL INFORMATION:
;  APPLICANT: Peak, Ian
;  APPLICANT:  Jennings, Michael
;  TITLE OF INVENTION: MODIFIED SURFACE ANTIGEN
;  FILE REFERENCE: 8795-24U1
;  CURRENT APPLICATION NUMBER: US/09/771,382
;  CURRENT FILING DATE:  2001-01-25
;  PRIOR APPLICATION NUMBER: US 60/177,917
;  PRIOR FILING DATE: 2000-01-25
;  NUMBER OF SEQ ID NOS: 52
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
;   LENGTH: 599
;   TYPE: PRT
;   ORGANISM: Neisseria meningitidis
US-09-771-382-6

  Query Match             96.6%;  Score 1881.5;  DB 10;  Length 599;
  Best Local Similarity   97.6%;  Pred. No. 4.1e-144;
  Matches  373;  Conservative    1;  Mismatches    7;  Indels    1;  Gaps    1;

Qy          2 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKT 61
              |||||||||||||||||||||||||||||| |||||||||||||||||||||||||:|
Db        218 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVHTYDTVEFLSADTKTTTVNVESKDNGKRT 277

Qy         62 EVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 121
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        278 EVKIGAKTSVIKEKDGKLVTGKGKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 337

Qy        122 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGW 181
              |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        338 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVKYDVNVGDALNVNQLQNSGW 397

Qy        182 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 241
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        398 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 457

Qy        242 SLGAGADAPTLSVDG-DALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNR 300
              |||||||||||||    ||||||| ||||||||||||||||||||||||||||||||||||
Db        458 SLGAGADAPTLSVDDKGALNVGSKDANKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNR 517

Qy        301 IDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWII 360
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        518 IDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWII 577

Qy        361 KGTASGNSRGHFGASASVGYQW 382
              ||||||||||||||||||||||
Db        578 KGTASGNSRGHFGASASVGYQW 599


Search completed: October  6, 2003, 10:19:51
Job time : 20.136 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: October 6, 2003, 09:13:35 ; Search time 10.9348 Seconds
(without alignments)
3359.577 Million cell updates/sec

Title: US-09-771-382-38
Perfect score: 1947
Sequence: 1 NRAASVKDVLNAGWNIKGVK.....TASGNSRGHFGASASVGYQW 382

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_76:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1941	99.7	591	2	G81133	adhesin NMB0992 [i
2	1854.5	95.2	592	2	A81888	probable surface f
3	336.5	17.3	2059	2	D82671	surface protein XF
4	335	17.2	1190	2	A82615	surface protein XF
5	325.5	16.7	1588	2	A86036	probable adhesin Z
6	325.5	16.7	1588	2	H91188	probable adhesin E
7	318	16.3	1107	2	AC0976	probable autotrans
8	312	16.0	658	2	AH0110	probable surface p
9	247	12.7	298	2	I64138	adhesin homolog HI
10	185	9.5	365	2	AB3486	cell surface prote
11	176.5	9.1	1004	2	C82672	surface-exposed ou
12	175	9.0	585	2	F90961	flagellin [importe
13	175	9.0	585	2	F85809	hypothetical prote
14	163.5	8.4	584	2	C48658	flagellin - Escher
15	163	8.4	364	2	A81019	adhesin/invasin, p
16	159	8.2	1536	2	A43855	high-molecular-wei
17	159	8.2	1910	2	AF0394	probable adhesin h
18	158.5	8.1	1477	2	B43855	high-molecular-wei
19	156.5	8.0	933	2	S41539	fibrinogen-binding
20	156	8.0	4919	2	T31105	hypothetical prote
21	155.5	8.0	989	2	D89852	fibrinogen-binding
22	155	8.0	1335	2	T17508	glycoprotein Vp260
23	154.5	7.9	2249	2	A41477	190K surface antig
24	154	7.9	364	2	AE0169	probable exported
25	152.5	7.8	595	2	A48658	flagellin - Escher
26	152.5	7.8	1091	2	G64964	hypothetical prote
27	152.5	7.8	1635	2	AI0452	hemolysin [importe
28	152.5	7.8	1655	2	E97835	hypothetical prote
29	152	7.8	2020	2	C48399	ABC-type transport
30	152	7.8	2273	2	T09083	hemagglutinin/hemo
31	151	7.8	980	2	H90681	probable flagellin
32	151	7.8	980	2	D85532	probable structura
33	150.5	7.7	1461	2	E90696	hypothetical prote
34	150	7.7	1300	2	S07575	outer membrane pro
35	150	7.7	5188	2	B85547	probable RTX famil
36	150	7.7	5291	2	F90696	hypothetical prote
37	148.5	7.6	595	2	B48658	flagellin - Escher
38	148.5	7.6	2551	2	B98047	hypothetical prote
39	148	7.6	1651	2	JC1340	outer membrane pro
40	147.5	7.6	338	2	D90697	adhesin/invasin-li
41	147.5	7.6	338	2	G85547	adhesin/invasin-li
42	147	7.6	936	2	I40711	sapB protein - Cam
43	147	7.6	1361	2	T03415	S-layer protein -
44	146.5	7.5	906	2	C85739	hypothetical prote
45	146.5	7.5	1011	2	F90879	probable BigA-like

ALIGNMENTS

RESULT 1
G81133
adhesin NMB0992 [imported] - Neisseria meningitidis (strain MC58 serogroup B)
C;Species: Neisseria meningitidis
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001
C;Accession: G81133
R;Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B. ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Masignani, V.; Pizza, M. Science 287, 1809-1815, 2000
A;Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.;
A;Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A;Reference number: A81000; MUID:20175755; PMID:10710307
A;Accession: G81133
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-591 <TET>
A;Cross-references: GB:AE002450; GB:AE002098; NID:g7226229; PIDN:AAF41395.1; PID:g722
A;Experimental source: serogroup B, strain MC58
C;Genetics:
A;Gene: NMB0992

Query Match 99.7%; Score 1941; DB 2; Length 591;
Best Local Similarity 100.0%; Pred. No. 3.7e-103;
Matches 381; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          2 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKT 61
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        211 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKT 270

Qy         62 EVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        271 EVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 330

Qy        122 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGW 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        331 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGW 390

Qy        182 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        391 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 450

Qy        242 SLGAGADAPTLSVDGDALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRI 301
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        451 SLGAGADAPTLSVDGDALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRI 510

Qy        302 DNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIK 361
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        511 DNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIK 570

Qy        362 GTASGNSRGHFGASASVGYQW 382
              |||||||||||||||||||||
```

```
Db        571 GTASGNSRGHFGASASVGYQW 591


RESULT 2
A81888
probable surface fibril protein NMA1200 [imported] - Neisseria meningitidis (strain Z249
C;Species: Neisseria meningitidis
C;Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001
C;Accession: A81888
R;Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel
; Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,
Nature 404, 502-506, 2000
A;Title: Complete DNA sequence of a serogroup A strain of Neisseria menigitidis Z2491.
A;Reference number: A81775; MUID:20222556; PMID:10761919
A;Accession: A81888
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-592 <PAR>
A;Cross-references: GB:AL162755; GB:AL157959; NID:g7379742; PIDN:CAB84461.1; PID:g737989
A;Experimental source: serogroup A, strain Z2491
C;Genetics:
A;Gene: NMA1200

  Query Match             95.2%;  Score 1854.5;  DB 2;  Length 592;
  Best Local Similarity   95.8%;  Pred. No. 3.1e-98;
  Matches  368;  Conservative    6;  Mismatches    7;  Indels    3;  Gaps    2;

Qy          2 RAASVKDVLNAGWNIKGVKPGTTA--SDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGK 59
              ||||:||||||||||||||| |:|   |:|||||||||||||||||||||||||||||||
Db        209 RAASIKDVLNAGWNIKGVKTGSTTGQSENVDFVRTYDTVEFLSADTKTTTVNVESKDNGK 268

Qy         60 KTEVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTT 119
              :|||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        269 RTEVKIGAKTSVIKEKDGKLVTGKGKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTT 328

Qy        120 ANGQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNS 179
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        329 ANGQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNS 388

Qy        180 GWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFS 239
              ||||||||||||||||||||||||||||||||||||||||||:|||||||||||| ||||
Db        389 GWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEISRNGKNIDIATSMAPQFS 448

Qy        240 SVSLGAGADAPTLSVDGD-ALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLN 298
              |||||||||||||||| : ||||||| ||||||||||||||||||||||||||||||||
Db        449 SVSLGAGADAPTLSVDDEGALNVGSKDANKPVRITNVAPGVKEGDVTNVAQLKGVAQNLN 508

Qy        299 NRIDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNW 358
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        509 NRIDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNW 568

Qy        359 IIKGTASGNSRGHFGASASVGYQW 382
              ||||||||||||||||||||||||
Db        569 IIKGTASGNSRGHFGASASVGYQW 592


RESULT 3
D82671
surface protein XF1529 [imported] - Xylella fastidiosa (strain 9a5c)
C;Species: Xylella fastidiosa
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C;Accession: D82671
R;anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000
A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A;Reference number: A82515; MUID:20365717; PMID:10910347
A;Note: for a complete list of authors see reference number A59328 below
A;Accession: D82671
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-2059 <SIM>
A;Cross-references: GB:AE003982; GB:AE003849; NID:g9106554; PIDN:AAF84338.1; GSPDB:GN001
A;Experimental source: strain 9a5c
R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer
as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Fr
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; La
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins
A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.
, F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri,
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawa
A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silv
M.; Tsuhako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.
A;Reference number: A59328
A;Contents: annotation
C;Genetics:
A;Gene: XF1529

  Query Match             17.3%;  Score 336.5;  DB 2;  Length 2059;
  Best Local Similarity   25.1%;  Pred. No. 2.7e-11;
  Matches  128;  Conservative   63;  Mismatches  154;  Indels  165;  Gaps   19;

Qy          5 SVKDVLNAGWNI--------KGVKPGTTASDNVD-------FVRTYDTVEFLSADTK--- 46
              || : :: || :         | |  ||    | |           : | |  || | |
Db       1583 SVSEAVDKGWTLTASGANGSKVVSGGTVDLKNTDGNLAISKSGDSNDVVFNLSKDFKVDE 1642

Qy         47 ----TTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKDKGEN-------GSSTDEG 95
                 | || :    |  ::| :||    |         :| : |:         | :  :
Db       1643 VTAGNTVVNTDGVKVG--SDVSLGAMGLFIANGPSVTASGFNAGDKVISHVAVGMADTDA 1700

Qy         96 EGLVTAKEVIDAVNKAGWRMKTTT----------ANGQTGQ---ADKFETVTSGTNV--- 139
              :   |: : :|     |  :|           :| ||    |    |  ||
Db       1701 VNVSQLKQAVQSVTVKATRYYSTNDGGTQGGNYDGDGATGSKAIAAGVGTQASGEGAAAV 1760

Qy        140 ---TFASGKGTTA-------------------------------------------TVS-- 152
                |||||:||                                           |||
Db       1761 GSGAAASGKGSTAIGRNAIASADGSVALGDGAKDGGRGAESYTGKYSGVQNNTVGTVSVG 1820

Qy        153 ---KDDQGNITVMYDVNVG-DALNVNQL----QNSGWNLDS------------------- 185
                | :  :|: : |     ||:|: ||    | |    |
Db       1821 DAAKGETRSISNVADAKEAMDAVNLRQLDAVAQKSNLQTDDMRHEINNIEDVFKITKGDS 1880

Qy        186 ---------KAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTP 236
                      |:|  :   :||  | : ||     |::| | : |  ||       |:
Db       1881 ASSVKGMGVNAMAIGTNAAVSGTESVALGK---NTNVSADNAVAI-GNG-------SVAD 1929

Qy        237 QFSSVSLGAGADAPTLSVDGDALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQN 296
              : :|||:|:|              ||::     ::|||| |  : |  ||:||
Db       1930 RANSVSVGSG--------------GSER-----QVTNVAAGTADTDAVNVSQLNQGLIT 1969

Qy        297 LNNRIDNVDGNAR----AGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSI 352
                | : || |     |:| ||||| | |||: |: | ::|  :|:|::  |:| |::
Db       1970 AKQYTDGMVGNLRRETSGGVAAAIATANLPQAYVQGRGMTSVGVSSYQGQSAIAVGVSAV 2029

Qy        353 SDGGNWIIKGTASGNSRGHFGASASVGYQW 382
              |: |:|: | : | |:| | |  | |||||
Db       2030 SESGHWVFKFSGSANTRSHVGVGAGVGYQW 2059


RESULT 4
A82615
surface protein XF1981 [imported] - Xylella fastidiosa (strain 9a5c)
C;Species: Xylella fastidiosa
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C;Accession: A82615
R;anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Seq
Nature 406, 151-157, 2000
A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A;Reference number: A82515; MUID:20365717; PMID:10910347
A;Note: for a complete list of authors see reference number A59328 below
A;Accession: A82615
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1190 <SIM>
```

```
A;Cross-references: GB:AE004017; GB:AE003849; NID:g9107083; PIDN:AAF84783.1; GSPDB:GN001
A;Experimental source: strain 9a5c
R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H
as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigr
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E
A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
, F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
M.; Tsuhako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A;Reference number: A59328
A;Contents: annotation
C;Genetics:
A;Gene: XF1981

  Query Match             17.2%;  Score 335;  DB 2;  Length 1190;
  Best Local Similarity   26.9%;  Pred. No. 1.7e-11;
  Matches  116;  Conservative   54;  Mismatches  151;  Indels  110;  Gaps   16;

Qy         10 LNAG-WNIKGVKPGTTASD-----------------NVDFVRTYDTVEFLSADTKTTTVN 51
              :|||   |  |  ||  :|                 :| :  |||          |:   |
Db        812 INAGSQKITNVAAGTADTDAVNLSQLNTAMAGSGAKSVHYYSTYD------GGTQGGNYN 865

Qy         52 VESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKDKGENG-SSTDEGEGLVTAKEVIDAVNK 110
              :    :  | :|   |    :|    |     :|  ||  |   | : :   |:
Db        866 GDGATGTRSIAVGVGTLASA----EGATAVGSGAAASGKGSTAIGRNAVASADGSVALGD 921

Qy        111 AGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDA 170
               | :     |   ||:    :  | ||     : || | |||     |:    :    ||
Db        922 -GAKDGARGAESYTGKYSGLQNNTVGTVSVGDASKGETRTVS-----NVADAKEAT--DA 973

Qy        171 LNVNQL----QNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGK 226
              :|: ||    |::   :|:|  : | |:                | :|: ||
Db        974 VNLRQLDRVAQDANRYVDNKIESLSEGQTF-------------VKVNSLNN-------- 1011

Qy        227 NIDIATSMTPQFSSVSLGAGADAPTLSV-----DGDALNVGSK----KDNKPV------- 270
                | ||      : || ||  : |       |:: :|:|     ||
Db       1012 ------SATP------IAAGVDATAIGVGATASGADSIAMGNKASASADNAVAIGNHSVA 1059

Qy        271 --------------RITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNAR----AG 311
                           ::|||| |  : |  ||:||           | | |: |     |
Db       1060 DRANTVSVGSAGSERQVTNVAAGTADTDAVNVSQLNQGLITAKQYTDGVVGSLRRDTDGG 1119

Qy        312 IAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGH 371
              :| ||||| | |||:||: | ::|  :|||::  |:| ||:|: | |: | : | |:|
Db       1120 VAAAIATANLPQAYIPGRGMTSVGVSSYRGQSAIAVGVSSVSESGRWVFKFSGSANTRSQ 1179

Qy        372 FGASASVGYQW 382
               |  | |||||
Db       1180 VGIGAGVGYQW 1190


RESULT 5
A86036
probable adhesin Z5029 [imported] - Escherichia coli (strain O157:H7, substrain EDL933)
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C;Accession: A86036
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: A86036
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1588 <STO>
A;Cross-references: GB:AE005174; NID:g12518349; PIDN:AAG58749.1; GSPDB:GN00145; UWGP:Z5C
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: Z5029

  Query Match             16.7%;  Score 325.5;  DB 2;  Length 1588;
  Best Local Similarity   26.2%;  Pred. No. 8.5e-11;
  Matches  117;  Conservative   61;  Mismatches  175;  Indels   94;  Gaps   17;

Qy          3 AASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVE-FLSADTKTTTVNVESKDNGKK- 60
              | :|: : ||         |  |:    :     | |  |:  | :  :  ::  || |
Db       1169 AVTVRQLQNA--------IGAVATTPTKYFHANSTEEDSLAVGTDSLAMGAKTIVNGDKG 1220

Qy         61 --------------TEVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVID 106
                           : ||:   ||      :  |     |||:|  |         :|
Db       1221 IGIGYGAYVDANALNGIAIGSNAQVIHV--NSIAIG-----NGSTTTRGAQTNYTAYNMD 1273

Qy        107 AVNKAGWRMKTTTANGQ-------TGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNI 159
              |   :       :|:||        | ||     |   ||    | | | ||:: | :|
Db       1274 APQNSVGEFSVGSADGQRQITNVAAGSAD-----TDAVNV--GQLKVTDAQVSQNTQ-SI 1325

Qy        160 TVM-------------YDVNVGDAL----------NVNQLQNSGWNLDSKAVAGSSGKVI 196
              | :               :  :|| :          | : :  |    || |:   |
Db       1326 TNLDNRVTNLDSRVTNIENGIGDIVTTGSTKYFKTNTDGVDASAQGKDSVAIGSGSIAAA 1385

Qy        197 SGNVSPSKGKM---DETVNINAGNNIEITRN---GKN----IDIATSMTPQFSSVSLGAG 246
               :|:   | :   : |::: :  |     |   |||    :::|   :  :   |
Db       1386 DNSVALGTGSVATEENTISVGSSTNQRRITNVAAGKNATDAVNVAQLKSSEAGGVRYDTK 1445

Qy        247 ADAPTLSVDGDALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQ----------- 295
              ||    |:|   : :|   :    ||:||: ||   || | ||||    |
Db       1446 ADG---SIDYSNITLGG-GNGGTTRISNVSAGVNNNDVVNYAQLKQSVQETKQYTDQRMV 1501

Qy        296 NLNNRIDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDG 355
               ::|::   :     ||| |:|  || ||| || || :|||||| ||:  |:| | :|
Db       1502 EMDNKLSKTESKLSGGIASAMAMTGLPQAYTPGASMASIGGGTYNGESAVALGVSMVSAN 1561

Qy        356 GNWIIKGTASGNSRGHFGASASVGYQW 382
              | |: |   | ||:| : |:   | ||
Db       1562 GRWVYKLQGSTNSQGEYSAALGAGIQW 1588


RESULT 6
H91188
probable adhesin ECs4480 [similarity] - Escherichia coli (strain O157:H7, substrain R
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 02-Nov-2001
C;Accession: H91188
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and g
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: H91188
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1588 <HAY>
A;Cross-references: GB:BA000007; PIDN:BAB37903.1; PID:g13363955; GSPDB:GN00154
A;Experimental source: strain O157:H7, substrain RIMD 0509952
C;Genetics:
A;Gene: ECs4480

  Query Match             16.7%;  Score 325.5;  DB 2;  Length 1588;
  Best Local Similarity   26.2%;  Pred. No. 8.5e-11;
  Matches  117;  Conservative   61;  Mismatches  175;  Indels   94;  Gaps   17;

Qy          3 AASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVE-FLSADTKTTTVNVESKDNGKK- 60
              | :|: : ||         |  |:    :     | |  |:  | :  :  ::  || |
Db       1169 AVTVRQLQNA--------IGAVATTPTKYFHANSTEEDSLAVGTDSLAMGAKTIVNGDKG 1220

Qy         61 --------------TEVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVID 106
                           : ||:   ||      :  |     |||:|  |         :|
Db       1221 IGIGYGAYVDANALNGIAIGSNAQVIHV--NSIAIG-----NGSTTTRGAQTNYTAYNMD 1273

Qy        107 AVNKAGWRMKTTTANGQ-------TGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNI 159
```

```
Db       1274 APQNSVGEFSVGSADGQRQITNVAAGSAD-----TDAVNV--GQLKVTDAQVSQNTQ-SI 1325

Qy        160 TVM-------------YDVNVGDAL----------NVNQLQNSGWNLDSKAVAGSSGKVI 196
Db       1326 TNLDNRVTNLDSRVTNIENGIGDIVTTGSTKYFKTNTDGVDASAQGKDSVAIGSGSIAAA 1385

Qy        197 SGNVSPSKGKM---DETVNINAGNNIEITRN---GKN----IDIATSMTPQFSSVSLGAG 246
Db       1386 DNSVALGTGSVATEENTISVGSSTNQRRITNVAAGKNATDAVNVAQLKSSEAGGVRYDTK 1445

Qy        247 ADAPTLSVDGDALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQ---------- 295
Db       1446 ADG---SIDYSNITLGG-GNGGTTRISNVSAGVNNNDVVNYAQLKQSVQETKQYTDQRMV 1501

Qy        296 NLNNRIDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDG 355
Db       1502 EMDNKLSKTESKLSGGIASAMAMTGLPQAYTPGASMASIGGGTYNGESAVALGVSMVSAN 1561

Qy        356 GNWIIKGTASGNSRGHFGASASVGYQW 382
Db       1562 GRWVYKLQGSTNSQGEYSAALGAGIQW 1588


RESULT 7
AC0976
probable autotransporter sapB [imported] - Salmonella enterica subsp. enterica serovar T
C;Species: Salmonella enterica subsp. enterica serovar Typhi
A;Note: this species has also been called Salmonella typhi
C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C;Accession: AC0976
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A;Reference number: AB0502; MUID:21534947; PMID:11677608
A;Accession: AC0976
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1107 <PAR>
A;Cross-references: GB:AL513382; PIDN:CAD03303.1; PID:g16504923; GSPDB:GN00176
C;Genetics:
A;Gene: sapB

  Query Match             16.3%;  Score 318;  DB 2;  Length 1107;
  Best Local Similarity   25.9%;  Pred. No. 1.5e-10;
  Matches  115;  Conservative   60;  Mismatches  180;  Indels   89;  Gaps   15;

Qy         16 IKGVKPGTTASDNVDFVRTYDTVEFL-----------SADTKTTTVNVESKDNGKKTEVK 64
Db        676 ITNVADGSEAQDAVTVRQLQNAIGAVTTTPTKYYHANSTEEDSLAVGTDSLAMGAKTIVN 735

Qy         65 ------IGAKTSVIKEKDGKLVTGKDKGE---------NGSSTDEG-EGLVTAKEVIDAV 108
Db        736 ADAGIGIGLNTLVMADAINGIAIGSNARANHANSIAMGNGSQTTRGAQTDYTAYNMDTPQ 795

Qy        109 NKAGWRMKTTTANGQ-------TGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITV 161
Db        796 NSVG-EFSVGSEDGQRQITNVAAGSAD-----TDAVNV--GQLKVTDAQVSRNTQSITNL 847

Qy        162 MYDVN------------VGDALNVNQLQNSGWNLD--------SKAVAGSSGKVISGNVS 201
Db        848 NTQVSNLDTRVTNIENGIGDIVTTGSTKYFKTNTDGADANAQGADSVAIGSGSIAAAENS 907

Qy        202 PSKGK---MDETVNINAGNNIEITR---------NGKNIDIATSMTPQFSSVSLGAGADA 249
Db        908 VALGTNSVADEANTVSVGSSTQQRRITNVAAGVNNTDAVNVAQLKASEAGSVRYETNADG 967

Qy        250 PTLSVDGDALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQ-----------NLN 298
Db        968 ---SVNYSVLNLGDGSGG-TTRIGNVSAAVNDTDAVNYAQLKRSVEEANTYTDQKMGEMN 1023

Qy        299 NRIDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNW 358
Db       1024 SKIKGIENKMSGGIASAMAMAGLPQAYAPGANMTSIAGGTFNGESAVAIGVSMVSESGGW 1083

Qy        359 IIKGTASGNSRGHFGASASVGYQW 382
Db       1084 VYKLQGTSNSQGDYSAAIGAGFQW 1107


RESULT 8
AH0110
probable surface protein (partial) YPO0902 [imported] - Yersinia pestis (strain CO92)
C;Species: Yersinia pestis
C;Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 02-Nov-2001
C;Accession: AH0110
R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrel
Nature 413, 523-527, 2001
A;Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A;Reference number: AB0001; MUID:21470413; PMID:11586360
A;Accession: AH0110
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-658 <KUR>
A;Cross-references: GB:AL590842; PIDN:CAC89747.1; PID:g15978974; GSPDB:GN00175
C;Genetics:
A;Gene: YPO0902

  Query Match             16.0%;  Score 312;  DB 2;  Length 658;
  Best Local Similarity   26.2%;  Pred. No. 1.7e-10;
  Matches  107;  Conservative   58;  Mismatches  151;  Indels   92;  Gaps   15;

Qy          5 SVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKT--- 61
Db        313 SMASIANGG-GVKYFHANSTQPDSV-----------ASGTNSVAIGPASLASGNAALAS 359

Qy         62 ---EVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTT 118
Db        360 GAGAVAIG--DGAAASADGSVAIGQGSGDNGRGVENYIG----------------KYSN 400

Qy        119 TANGQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQN 178
Db        401 ASNTSSG------TVSVGNTAT-----GETRTVSNVADG-------LQATDAVNLRQLDG 442

Qy        179 SGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNI-EITRNGKNIDI--ATSMT 235
Db        443 ----------IAASIVVVENNVSGLQNGTDGMFQVNNSSGLAKPSATGANSATGGAGSVA 492

Qy        236 PQFSSVSLGAGADAP----------TLSVDGDALNVGSKKDNKPVRITNVAPGVKEGDVT 285
Db        493 SGNNSTAFGSGAKATAANSAALGANSVADRANSVSVGSVGNER--QITNVAPATQGTDAV 550

Qy        286 NVAQLKGVAQN----LNNRIDNVDGNAR-------AGIAQAIATAGLVQAYLPGKSMMAI 334
Db        551 NFDQLKSISNQTNAYTNQRYSELKQDLRKQNSVLSAGIASAMSMASLTQPYTSGSSMTTI 610

Qy        335 GGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNSRGHFGASASVGYQW 382
Db        611 GAASYRGQSALSLGVSSISDSGRWVSKLQASSNTQGDFGIGVGVGYQW 658


RESULT 9
I64138
adhesin homolog HI1732 - Haemophilus influenzae (strain Rd KW20)
C;Species: Haemophilus influenzae
C;Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 24-Oct-1997
C;Accession: I64138
R;Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage
; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman
, D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M.
Science 269, 496-512, 1995
```

```
A;Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,
A;Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
A;Reference number: A64000; MUID:95350630; PMID:7542800
A;Accession: I64138
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-298 <TIGR>
A;Cross-references: GB:U32846; GB:L42023; NID:g1574588; PID:g1574589; TIGR:HI1732

  Query Match             12.7%;  Score 247;  DB 2;  Length 298;
  Best Local Similarity   52.6%;  Pred. No. 3.2e-07;
  Matches   51;  Conservative   15;  Mismatches   29;  Indels    2;  Gaps    1;

Qy          2 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKT 61
              |||::|||||||||||| |     ::||| |  || |||:: |  |  | : :|:||| |
Db        195 RAATIKDVLNAGWNIKGAKVAGGNTENVDLVAGYDNVEFITGDKNTLDVVLTAKENGKTT 254

Qy         62 EVKIGAKTSVIKEKDGKLVTGKD--KGENGSSTDEGE 96
              |||   ||||||: :|||:|||       |::|:  |
Db        255 EVKFTPKTSVIKDNNGKLLTGKQLKDANTGTATNATE 291


RESULT 10
AB3486
cell surface protein [imported] - Brucella melitensis (strain 16M)
C;Species: Brucella melitensis
C;Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 01-Feb-2002
C;Accession: AB3486
R;DelVecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,
.; Mazur, M.; Goltsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letess
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A;Title: The genome sequence of the facultative intracellular pathogen Brucella melitens
A;Reference number: AD3252; PMID:11756688
A;Accession: AB3486
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-365 <KUR>
A;Cross-references: GB:AE008917; PIDN:AAL53053.1; PID:g17983913; GSPDB:GN00190
A;Experimental source: strain 16M
C;Genetics:
A;Gene: BMEI1872
A;Map position: I

  Query Match              9.5%;  Score 185;  DB 2;  Length 365;
  Best Local Similarity   24.3%;  Pred. No. 0.0014;
  Matches  102;  Conservative   43;  Mismatches  113;  Indels  162;  Gaps   20;

Qy          4 ASVKDVLN---AGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGK- 59
              |:: |:|:   :|  ||      :  :|:             :  | :  |  :|  :|:
Db         32 ANINDILHDIESGGGIKYFHANSIGADS-----------RALGTNSIAVGSDSVASGEG 79

Qy         60 KTEVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDE-----GEGLVTAKEVIDAVNKAGWR 114
                |  ||: |     | :     ||| :: |       | |  |:: |
Db         80 SISVGNGAQASA----HGSVAL----GENAAAPDANSVALGAGSKTSEVV---------A 122

Qy        115 MKTTTANGQ----TGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDA 170
               | || |||     | |       |||       ||   |::    | |:    |   ||
Db        123 TKGTTINGQYYDFAGDA------PSGT--VSVGDKGAERTITNVAAGRIS----VESTDA 170

Qy        171 LNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDI 230
              :| :||     |  ::|:                    |: ||    :| | |
Db        171 VNGSQL-----NAVNQAIE---------------------NLAAG----VTENDK---- 195

Qy        231 ATSMTPQFSSVSLGAGADAPTLSVDGDALNVGSKKDN---------KPVRITNVAPGVKE 281
                                  ||  |  : |:||::          || : ||| || :
Db        196 --------------------FSVKYDRHSDGTKKNSMTLQGWDSATPVVLANVADGVHK 234

Qy        282 GDVTNVAQLK-GVAQNLNNRIDNVDGNARAGIAQA----------------------I 316
               |  ||:||| |::  |       |  |   : ||
Db        235 NDAVNVSQLKAGLSTTLGEAKAYTDQTALQTLDQANAYTDKKFGKLNEDIVATRIEARQA 294

Qy        317 ATAGLVQAYL-----PGKSMMAIGGGTYRGEAGYAIG-------------YSSISDGGNW 358
              |  ||  | |     |||   ||||| :|||   |:|              |: : ||||

Db        295 AAIGLAAASLRYDDRPGKISAAIGGGFWRGEGAVALGLGHTSEDQRMRSNLSAATSGGNW 354


RESULT 11
C82672
surface-exposed outer membrane protein XF1516 [imported] - Xylella fastidiosa (strain
C;Species: Xylella fastidiosa
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C;Accession: C82672
R;anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Seq
Nature 406, 151-157, 2000
A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A;Reference number: A82515; MUID:20365717; PMID:10910347
A;Note: for a complete list of authors see reference number A59328 below
A;Accession: C82672
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1004 <SIM>
A;Cross-references: GB:AE003981; GB:AE003849; NID:g9106543; PIDN:AAF84325.1; GSPDB:GN
A;Experimental source: strain 9a5c
R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer
as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Fr
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; La
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins
A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.
, F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri,
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawa
A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silv
M.; Tsuhako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.
A;Reference number: A59328
A;Contents: annotation
C;Genetics:
A;Gene: XF1516

  Query Match              9.1%;  Score 176.5;  DB 2;  Length 1004;
  Best Local Similarity   25.8%;  Pred. No. 0.014;
  Matches   98;  Conservative   58;  Mismatches  141;  Indels   83;  Gaps   18;

Qy         53 ESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKDKGE-NGSSTDEGEGLV--TAKEVI---- 105
              :| :||    | :|: ::|       |  |     | |  |   |:|:     : ::
Db        139 QSANNG----VALGSNSTVSGVNSVALGAGSMASELNVISVGGGDGVTGPAVRRIVNVGD 194

Qy        106 -----DAVNKAGWRMKTTTANGQTGQADKFETV-----TSGTNVTFASGKGTTATVS-KD 154
                    |||||:     | : |      |   :|:      :|::|  |||| :|| |
Db        195 GIGNNDAVNKSQLDGVTASVN---DVAASVKTIALTNQVTGSSVASASGKESTAIGSGAQ 251

Qy        155 DQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNIN 214
                 : || :    |      | :  |    || |   :|  | : :||  :| :    |
Db        252 AVADNTVAF----GGRAIANAVGASALGFDSHAKGINSTTVGTQSVSLGQGGVSLGYNSF 307

Qy        215 AG----NNIEITRNG----KNIDIATSMTPQFSSVSLGAGADAPTLSV----DGDALNVG 262
              |    | : :  |      : :|           ||:||:|: |   :|     || |
Db        308 VGEGSFNGLALGSNSLVLLQGVD----------SVALGSGSMASEPNVVSVGSGDGLRGP 357

Qy        263 SKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNV---------DGNARAGIA 313
              : :     || ||  |:   |  | :|| ||  ::|: : :|          |:  | ::
Db        358 AVR-----RIVNVGDGIGNNDAVNKSQLDGVTASVNDVVASVKNIAGAIQITGSGVASVS 412

Qy        314 QAIATAGLVQAYLPGKSMMAIG--------GGTYRGEAGYAIGYSSISDGGNWIIKGTA- 364
                :||    |   | | :|:|        | :  |  |:|:| :| : ||    : ||
Db        413 GQDSTAAGASAQAAGDSSIALGARSRANAIGSSALGVDGHALGANSTALGG----QSTAI 468

Qy        365 --SGNSRGH---FGASASVG 379
                | | |:    | ||: |
Db        469 SEGGTSLGYNSFVGQSATNG 488


RESULT 12
F90961
flagellin [imported] - Escherichia coli (strain O157:H7, substrain RIMD 0509952)
```

```
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001
C;Accession: F90961
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genc
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: F90961
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-585 <HAY>
A;Cross-references: GB:BA000007; PIDN:BAB36085.1; PID:g13362130; GSPDB:GN00154
A;Experimental source: strain O157:H7, substrain RIMD 0509952
C;Genetics:
A;Gene: ECs2662
C;Superfamily: flagellin

  Query Match             9.0%;  Score 175;  DB 2;  Length 585;
  Best Local Similarity   22.7%;  Pred. No. 0.0091;
  Matches   88;  Conservative   59;  Mismatches  167;  Indels   74;  Gaps   17;

Qy         22 GTTASDNVDFVR-----TYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKD 76
              || :  ::| ::       | :: :|  |:   ||| :||   |  :::||        |
Db        102 GTNSDSDLDSIQDEIKSRLDEIDRVSGQTQFNGVNVLAKDGSMK--IQVGANDGETITID 159

Qy         77 GKLVTGKDKGENGSSTDEGEGLVTAK-EVIDAVNKAGWRMKTTTANGQTGQAD-KFETVT 134
                | :     | || : : |:| :| |   :  :  || :: ||     ||  | | |
Db        160 LKKIDSDTLGLNGFNVN-GKGTITNKAATVSDLTSAGAKLNTT-----TGLYDLKTENTL 213

Qy        135 SGTNVTFAS-GKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSS- 192
              |:  |   | |   ||     | :   |:   ||         :|  :|: | |  |
Db        214 LTTDAAFDKLGNGDKVTV-----GGVDYTYNAKSGDFTTTK--STAGTGVDAAAQAADSA 266

Qy        193 --------------GKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQF 238
                           || ::|: :   | :  :   ::  || |  :   :| | ::|
Db        267 SKRDALAATLHADVGKSVNGSYTTKDGTV--SFETDSAGNITIGGSQAYVDDAGNLTTNN 324

Qy        239 SSVSLGA-----------GADAPTLSVDGDALNVGSKKDNKPVRITNVAPGVKEGDVTNV 287
              :  :  |           |:|  :|: :|    :        |   | ||| :  |
Db        325 AGSAAKADMKALLKAASEGSDGASLTFNGTEYTIAKA---TPATTTPVAPLIPGGITYQA 381

Qy        288 AQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQ---AYLPGKSMMAI-------GGG 337
                | |  :         :  | |  :      :|::     :  |:|  |         ||
Db        382 TVSKDVVLS--------ETKAAAATSSITFNSGVLSKTIGFTAGESSDAAKSYVDDKGGI 433

Qy        338 TYRGEAGYAIGYSSISDGGNWIIKGTAS 365
              |    | | : ||   | |:  : | ||
Db        434 T--NVADYTVSYSVNKDNGSVTVAGYAS 459


RESULT 13
F85809
hypothetical protein fliC [imported] - Escherichia coli (strain O157:H7, substrain EDL93
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C;Accession: F85809
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: F85809
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-585 <STO>
A;Cross-references: GB:AE005174; NID:g12516024; PIDN:AAG56938.1; GSPDB:GN00145; UWGP:Z3C
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: fliC
C;Superfamily: flagellin

  Query Match             9.0%;  Score 175;  DB 2;  Length 585;
  Best Local Similarity   22.7%;  Pred. No. 0.0091;
  Matches   88;  Conservative   59;  Mismatches  167;  Indels   74;  Gaps   17;

Qy         22 GTTASDNVDFVR-----TYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKD 76
              || :  ::| ::       | :: :|  |:   ||| :||   |  :::||        |
Db        102 GTNSDSDLDSIQDEIKSRLDEIDRVSGQTQFNGVNVLAKDGSMK--IQVGANDGETITID 159

Qy         77 GKLVTGKDKGENGSSTDEGEGLVTAK-EVIDAVNKAGWRMKTTTANGQTGQAD-KFETVT 134
                | :     | || : : |:| :| |   :  :  || :: ||     ||  | | |
Db        160 LKKIDSDTLGLNGFNVN-GKGTITNKAATVSDLTSAGAKLNTT-----TGLYDLKTENTL 213

Qy        135 SGTNVTFAS-GKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSS- 192
              |:  |   | |   ||     | :   |:   ||         :|  :|: | |  |
Db        214 LTTDAAFDKLGNGDKVTV-----GGVDYTYNAKSGDFTTTK--STAGTGVDAAAQAADSA 266

Qy        193 --------------GKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQF 238
                           || ::|: :   | :  :   ::  || |  :   :| | ::|
Db        267 SKRDALAATLHADVGKSVNGSYTTKDGTV--SFETDSAGNITIGGSQAYVDDAGNLTTNN 324

Qy        239 SSVSLGA-----------GADAPTLSVDGDALNVGSKKDNKPVRITNVAPGVKEGDVTNV 287
              :  :  |           |:|  :|: :|    :        |   | ||| :  |
Db        325 AGSAAKADMKALLKAASEGSDGASLTFNGTEYTIAKA---TPATTTPVAPLIPGGITYQA 381

Qy        288 AQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQ---AYLPGKSMMAI-------GGG 337
                | |  :         :  | |  :      :|::     :  |:|  |         ||
Db        382 TVSKDVVLS--------ETKAAAATSSITFNSGVLSKTIGFTAGESSDAAKSYVDDKGGI 433

Qy        338 TYRGEAGYAIGYSSISDGGNWIIKGTAS 365
              |    | | : ||   | |:  : | ||
Db        434 T--NVADYTVSYSVNKDNGSVTVAGYAS 459


RESULT 14
C48658
flagellin - Escherichia coli (strain U5-41)
C;Species: Escherichia coli
C;Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 26-Aug-1999
C;Accession: C48658
R;Schoenhals, G.; Whitfield, C.
J. Bacteriol. 175, 5395-5402, 1993
A;Title: Comparative analysis of flagellin sequences from Escherichia coli strains po
A;Reference number: A48658; MUID:93374833; PMID:8366026
A;Accession: C48658
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-584 <SCH>
A;Cross-references: GB:L07388; NID:g290438; PIDN:AAA23798.1; PID:g290439
C;Superfamily: flagellin

  Query Match             8.4%;  Score 163.5;  DB 2;  Length 584;
  Best Local Similarity   23.4%;  Pred. No. 0.041;
  Matches   91;  Conservative   57;  Mismatches  164;  Indels   77;  Gaps   19;

Qy         22 GTTASDNVDFVR-----TYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKD 76
              || :  ::| ::       | :: :|  |:   ||| :||   |  :::||        |
Db        102 GTNSDSDLDSIQDEIKSRLDEIDRVSGQTQFNGVNVLAKDGSMK--IQVGANDGETITID 159

Qy         77 GKLVTGKDKGENGSSTDEGEGLVTAK-EVIDAVNKAGWRMKTTTANGQTGQADKFETVTS 135
                | :     | || : : |:| :| |   :  :  || :: |||  |  |   | |
Db        160 LKKIDSDTLGLNGFNVN-GKGTITNKAATVSDLTSAGAKLNTTT--GLYGL--KTENTLL 214

Qy        136 GTNVTFAS-GKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSS-- 192
              |:  |   | |   ||   |      :|:   ||         :|  :|: | |  |
Db        215 TTDAAFDKLGNGDKVTVGGVD-----IYNAKSGDFTTTK--STAGTGVDAAAQATDSAK 266

Qy        193 ------------GKVISGNVSPSKGKMDETVNINAGNNIEITRNGKN--IDIATSMTPQ 237
                         || ::|: :   | :  |  :|||   ||  |    :| | ::|
Db        267 KRDALAATLHADVGKSVNGSYTTKDGTV-SFVTDSAGN---ITIGGSQAYVDDAGNLTTN 322

Qy        238 FSSVSLGA-----------GADAPTLSVDGDALNVGSKKDNKPVRITNVAPGVKEGDVTN 286
              :  :  |           |:|  :|: :|    :        |   : ||| :  |
Db        323 NAGSARKADMKALLKAASEGSDGASLTFNGTEYTIAKA---TPATTSPVAPLIPGGITYQ 379
```

```
Qy        287 VAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQ---AYLPGKSMMAI-------GG 336
Db        380 ATVSKDVVLS--------ETKAAAATSSVTFNSGVLSKTIGFTAGESSDAAKSYVDDKGG 431

Qy        337 GTYRGEAGYAIGYSSISDGGNWIIKGTAS 365
Db        432 IT--NVADYTVSYSVNKDNGSVTVAGYAS 458


RESULT 15
A81019
adhesin/invasin, probable NMB1994 [imported] - Neisseria meningitidis (strain MC58 serog
C;Species: Neisseria meningitidis
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001
C;Accession: A81019
R;Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;
ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Masignani, V.; Pizza, M.
Science 287, 1809-1815, 2000
A;Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ve
A;Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A;Reference number: A81000; MUID:20175755; PMID:10710307
A;Accession: A81019
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-364 <TET>
A;Cross-references: GB:AE002548; GB:AE002098; NID:g7227249; PIDN:AAF42321.1; PID:g722725
A;Experimental source: serogroup B, strain MC58
C;Genetics:
A;Gene: NMB1994

  Query Match             8.4%;  Score 163;  DB 2;  Length 364;
  Best Local Similarity   25.1%;  Pred. No. 0.025;
  Matches  103;  Conservative   54;  Mismatches  145;  Indels  108;  Gaps   23;

Qy          2 RAASVKDV--LNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTK-TTTVNVESKD-N 57
Db         34 KAATVAIVAAYNNGQEINGFKAGET---------IYDIGEDGTITQKDATAADVEADDFK 84

Qy         58 GKKTEVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAG----W 113
Db         85 G------LGLK---------KVVTNLTKTVNEN----------KQNVDAKVKAAESEIE 118

Qy        114 RMKTTTANGQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNV 173
Db        119 KLTTKLADTDAALADTDAALDETTNALNKLGENIT-TFAEETKTNI-----VKIDEKL-- 170

Qy        174 NQLQNSGWNLDSKAVAGSSGK------VISGNVSPSKGKMDETVNI--NAGNNIEITRNG 225
Db        171 -----------EAVADTVDKHAEAFNDIADSLDETNTKADEAVKTANEAKQTAEETK-- 216

Qy        226 KNIDIATSMTPQFSSVSLGAGADAPTLSVDGDALNVGSKKDNKPVRITNVAPGVKEGDVT 285
Db        217 QNVD---------AKVKAAETAAGKAEAAAGTANTAADKAEAVAAKVTD----IKADIAT 263

Qy        286 NVAQLKGVAQNLNNRIDNVDGN-------ARAGIAQAIATAGLVQAYLPGKSMMAIGGGT 338
Db        264 NKAD---IAKN-SARIDSLDKNVANLRKETRQGLAEQAALSGLFQPYNVGRFNVTAAVGG 319

Qy        339 YRGEAGYAIGYSSISDGGNWIIK-----GTASGNSRG-HFGASASVGYQW 382
Db        320 YKSESAVAIG-TGFRFTENFAAKAGVAVGTSSGSSAAYHVG----VNYEW 364


Search completed: October  6, 2003, 09:33:40
Job time : 11.9348 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: October 6, 2003, 09:06:20 ; Search time 5.89906 Seconds
(without alignments)
3045.266 Million cell updates/sec

Title: US-09-771-382-38
Perfect score: 1947
Sequence: 1 NRAASVKDVLNAGWNIKGVK.....TASGNSRGHFGASASVGYQW 382

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	154.5	7.9	2249	1	OMPA_RICRI	P15921 rickettsia
2	152.5	7.8	1039	1	AG43_ECOLI	P39180 escherichia
3	152.5	7.8	1655	1	OMPB_RICCN	Q9kka3 r outer mem
4	152	7.8	2003	1	YDBA_ECOLI	P33666 escherichia
5	150	7.7	1300	1	120K_RICRI	P14914 rickettsia
6	150	7.7	1654	1	OMPB_RICRI	Q53047 r outer mem
7	144	7.4	550	1	FLIC_SHIFL	Q08860 shigella fl
8	142	7.3	1953	1	BIGA_SALTY	P25927 salmonella
9	139.5	7.2	737	1	ALYS_ENTFA	P37710 enterococcu
10	139	7.1	1608	1	HLYA_SERMA	P15320 serratia ma
11	137.5	7.1	1007	1	Y741_CHLMU	Q9pjt6 chlamydia m
12	137	7.0	497	1	FLIC_ECOLI	P04949 escherichia
13	137	7.0	1567	1	ICEN_XANCT	P18127 xanthomonas
14	136.5	7.0	930	1	PMP8_CHLPN	Q9z393 chlamydia p
15	136	7.0	1577	1	HLYA_PROMI	P16466 proteus mir
16	136	7.0	1656	1	OMPB_RICJA	O06653 r outer mem
17	134.5	6.9	1286	1	AIDA_ECOLI	Q03155 escherichia
18	134.5	6.9	1569	1	YPJA_ECOLI	P52143 escherichia
19	133	6.8	928	1	PM10_CHLPN	Q9rb65 chlamydia p
20	133	6.8	1645	1	OMPB_RICTY	P96989 r outer mem
21	133	6.8	2021	1	OMPA_RICCN	Q52657 rickettsia
22	131	6.7	434	1	YADA_YERPS	P10858 yersinia ps
23	131	6.7	917	1	HXA3_HAEIN	P45355 haemophilus
24	131	6.7	1148	1	ICEK_PSESX	O30611 pseudomonas
25	131	6.7	1153	1	PVDB_PLAKN	P50493 plasmodium
26	131	6.7	1398	1	TOP2_PLAFK	P41001 plasmodium
27	130	6.7	2660	1	YEEJ_ECO57	Q8x8v7 escherichia
28	129.5	6.7	716	1	CX80_EUPOC	Q9n9x3 euplotes oc
29	129	6.6	1196	1	ICEV_PSESX	O33479 pseudomonas
30	129	6.6	3178	1	YS89_CAEEL	Q09624 caenorhabdi
31	128	6.6	575	1	FLA2_CAMJE	P22251 campylobact
32	128	6.6	933	1	SLAP_CAMFE	P35827 campylobact
33	128	6.6	1025	1	SLAP_CAUCR	P35828 caulobacter
34	127	6.5	817	1	YG4A_YEAST	P46949 saccharomyc
35	127	6.5	928	1	HXA2_HAEIN	P45354 haemophilus
36	126.5	6.5	1258	1	ICEN_ERWHE	P16239 erwinia her
37	126	6.5	1076	1	NUP1_YEAST	P20676 saccharomyc
38	126	6.5	1200	1	ICEN_PSESY	P06620 pseudomonas
39	126	6.5	1210	1	ICEN_PSEFL	P09815 pseudomonas
40	125.5	6.4	760	1	AMY_CLOAB	P23671 clostridium
41	125.5	6.4	1288	1	VACA_HELPJ	Q9zkw5 helicobacte
42	125	6.4	918	1	YMJB_CAEEL	P34487 caenorhabdi
43	124.5	6.4	365	1	FLC1_PROMI	P42272 proteus mir
44	124.5	6.4	455	1	YADA_YEREN	P31489 yersinia en
45	124	6.4	575	1	FLB2_CAMJE	P22252 campylobact

ALIGNMENTS

RESULT 1
OMPA_RICRI
ID OMPA_RICRI STANDARD; PRT; 2249 AA.
AC P15921;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Outer membrane protein A precursor (190 kDa antigen) (Cell surface
DE antigen) (rOmpA) (rOmp A).
GN OMPA.
OS Rickettsia rickettsii.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Rickettsiaceae; Rickettsieae; Rickettsia.
OX NCBI_TaxID=783;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=R;
RX MEDLINE=90354033; PubMed=2117568;
RA Anderson B.E., McDonald G.A., Jones D.C., Regnery R.L.;
RT "A protective protein antigen of Rickettsia rickettsii has tandemly
RT repeated, near-identical sequences.";
RL Infect. Immun. 58:2760-2769(1990).
CC -!- FUNCTION: ELICITS PROTECTIVE IMMUNITY.
CC -!- SUBCELLULAR LOCATION: CELL WALL. THIS BACTERIUM IS COVERED BY A
CC S-LAYER WITH HEXAGONAL SYMMETRY.
CC -!- PTM: GLYCOSYLATED (PROBABLE).
CC -!- SIMILARITY: BELONGS TO THE RICKETTSIAE OMPA/OMPB FAMILY.
CC ----------

CC ----------
DR EMBL; M31227; AAA26380.1; -.
DR PIR; A41477; A41477.
DR InterPro; IPR006315; Autotransport.
DR InterPro; IPR005546; Autotransporter.
DR Pfam; PF03797; Autotransporter; 1.
DR TIGRFAMs; TIGR01414; autotrans_barl; 3.
KW Antigen; Repeat; Signal; Cell wall; S-layer; Glycoprotein.
FT SIGNAL 1 28 POTENTIAL.
FT CHAIN 29 2249 OUTER MEMBRANE PROTEIN A.
FT DOMAIN 212 1180 13 X APPROXIMATE TANDEM REPEATS.
FT REPEAT 212 286 A (TYPE I).
FT REPEAT 287 358 B (TYPE II).
FT REPEAT 359 430 C (TYPE II).
FT REPEAT 431 505 D (TYPE I).
FT REPEAT 506 577 E (TYPE II).
FT REPEAT 578 652 F (TYPE I).
FT REPEAT 653 724 G (TYPE II).
FT REPEAT 725 799 H (TYPE I).
FT REPEAT 800 874 I (TYPE I).
FT REPEAT 875 949 J (TYPE I).

```
FT   REPEAT      950   1021       K (TYPE II).
FT   REPEAT     1022   1093       L (TYPE II).
FT   REPEAT     1094   1165       M (TYPE II).
FT   REPEAT     1166   1180       TYPE I (INCOMPLETE).
SQ   SEQUENCE   2249 AA;  224333 MW;  A9D6646C089DF087 CRC64;

  Query Match             7.9%;  Score 154.5;  DB 1;  Length 2249;
  Best Local Similarity   24.8%;  Pred. No. 0.44;
  Matches  119;  Conservative   41;  Mismatches  171;  Indels  149;  Gaps   25;

Qy          3 AASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTE 62
              ||||  : ||   : |    ||  |||  :     :  :: |   |        |   |
Db        833 AASVLTLTNANAVLTGAVDNTTGGDNVGVLNLNGALSQVTGDIGNT--------NSLAT- 883

Qy         63 VKIGAKT-----SVIKEKDGKL---------------------------------- 79
              : :|| |     :|||    ||
Db        884 ISVGAGTATLGGAVIKATTTKLTNAASVLTLTNANAVLTGAIDNTTGGDNVGVLNLNGAL 943

Qy         80 --VTGKDKGENGSST-DEGEGLVT-----AKEVIDAVNKAGWRMKTTTANGQTGQADKFE 131
                |||     |  :|   | |  |      |    :  |   :| |     ||  |
Db        944 SQVTGDIGNTNSLATISVGAGTATLGGAVIKATTTKLTDAASAVKFTNPVVVTGAIDNTG 1003

Qy        132 TVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGS 191
              :|  |||       |  |      ||   :  ||||  |   |:|      ::  :  :
Db       1004 NANNGI-VTFTGNSTVTGNV-----GNTNALATVNVGAGL--LQVQGGVVKANTINLTDN 1055

Qy        192 SGKVISGNVSPSKGKMDETVNINAG-----NNIEITRNGKNIDIATSMTPQFSSVSLGAG 246
              :  |   |      | :| | | | |      |  :| |  |       |   ::|::|||
Db       1056 ASAVTFTNPVVVTGAIDNTGNANNGIVTFTGNSTVTGNVGN-------TNALATVNVGAG 1108

Qy        247 ADAPTLSVDGDAL--NVGSKKDN-KPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDN 303
                | |  :  |  :  ||   |  ||  | |  | : |          | || |
Db       1109 ----LLQVQGGVVKANTINLTDNASAVTFTN--PVVVTGAIDNTG-------NANNGIVT 1155

Qy        304 VDGNARA----GIAQAIAT----AGL-VQAYLPGKSMMA----------------IGGGT 338
                ||:       |   |:||    ||: :||   | |: |                 : ||
Db       1156 FTGNSTVTGDIGNTNALATVNVGAGITLQA---GGSLAANNIDFGARSTLEFNGPLDGG- 1211

Qy        339 YRGEAGYAIGY---SSISDGGNWIIK-------------GTASGNSRGH---FGASASVG 379
                | || |     :|::| | |:              || :  : |    |   ||||
Db       1212 -----GKAIPYYFKGAIANGNNAILNVNTKLLTASHLTIGTVAEINIGAGNLFTIDASVG 1266


RESULT 2
AG43_ECOLI
ID   AG43_ECOLI     STANDARD;      PRT;  1039 AA.
AC   P39180; P75614; P76360; P97241; Q46771;
DT   01-FEB-1995 (Rel. 31, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Antigen 43 precursor (AG43) (Fluffing protein).
GN   FLU OR B2000.
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12 / MG1655;
RX   MEDLINE=97426617; PubMed=9278503;
RA   Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA   Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA   Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA   Mau B., Shao Y.;
RT   "The complete genome sequence of Escherichia coli K-12.";
RL   Science 277:1453-1474(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12;
RX   MEDLINE=97251358; PubMed=9097040;
RA   Itoh T., Aiba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K.,
RA   Kasai H., Kimura S., Kitakawa M., Kitagawa M., Makino K., Miki T.,
RA   Mizobuchi K., Mori H., Mori T., Motomura K., Nakade S., Nakamura Y.,
RA   Nashimoto H., Nishio Y., Oshima T., Saito N., Sampei G., Seki Y.,
RA   Sivasundaram S., Tagami H., Takeda J., Takemoto K., Wada C.,
RA   Yamamoto Y., Horiuchi T.;
RT   "A 460-kb DNA sequence of the Escherichia coli K-12 genome
RT   corresponding to the 40.1-50.0 min region on the linkage map.";
RL   DNA Res. 3:379-392(1996).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ML 308-225;
RA   Henderson I.R., Owen P.;
RL   Submitted (JAN-1997) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   PRELIMINARY SEQUENCE OF 53-78.
RC   STRAIN=ML 308-225;
RX   MEDLINE=89291704; PubMed=2661530;
RA   Caffrey P., Owen P.;
RT   "Purification and N-terminal sequence of the alpha subunit of antigen
RT   43, a unique protein complex associated with the outer membrane of
RT   Escherichia coli.";
RL   J. Bacteriol. 171:3634-3640(1989).
RN   [5]
RP   SEQUENCE OF 53-63.
RC   STRAIN=K12 / EMG2;
RX   MEDLINE=97443975; PubMed=9298646;
RA   Link A.J., Robison K., Church G.M.;
RT   "Comparing the predicted and observed properties of proteins encoded
RT   in the genome of Escherichia coli K-12.";
RL   Electrophoresis 18:1259-1313(1997).
RN   [6]
RP   GENE NAME.
RX   MEDLINE=97257509; PubMed=9103983;
RA   Henderson I.R., Meehan M., Owen P.;
RT   "Antigen 43, a phase-variable bipartite outer membrane protein,
RT   determines colony morphology and autoaggregation in Escherichia coli
RT   K-12.";
RL   FEMS Microbiol. Lett. 149:115-120(1997).
CC   -!- FUNCTION: CONTROLS COLONY FORM VARIATION AND AUTOAGGREGATION. MAY
CC       FUNCTION AS AN ADHESIN.
CC   -!- SUBUNIT: HETERODIMER OF THE PERIPHERAL MEMBRANE PROTEIN (ALPHA
CC       CHAIN) ANCHORED TO THE INTEGRAL OUTER MEMBRANE PROTEIN (BETA
CC       CHAIN).
CC   -!- SUBCELLULAR LOCATION: Outer membrane-associated.
CC   -!- SIMILARITY: TO ADHESIN AIDI-I AND TO BORDETELLA PERTACTIN.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE000291; AAC75061.1; ALT_INIT.
DR   EMBL; D90838; BAA15825.1; ALT_INIT.
DR   EMBL; D90839; BAA15832.1; ALT_INIT.
DR   EMBL; U24429; AAB47869.1; -.
DR   HSSP; P07505; 1SRD.
DR   EcoGene; EG12686; flu.
DR   InterPro; IPR006315; Autotransport.
DR   InterPro; IPR005546; Autotransporter.
DR   InterPro; IPR004899; Pertactin.
DR   Pfam; PF03797; Autotransporter; 1.
DR   Pfam; PF03212; Pertactin; 1.
DR   TIGRFAMs; TIGR01414; autotrans_barl; 1.
KW   Outer membrane; Signal; Complete proteome.
FT   SIGNAL        1     52
FT   CHAIN        53    551       ANTIGEN 43 ALPHA CHAIN.
FT   CHAIN       552   1039       ANTIGEN 43 BETA CHAIN.
FT   VARIANT       2      2       K -> N (IN STRAIN ML 308-225).
FT   VARIANT      41     42       SL -> FF (IN STRAIN ML 308-225).
FT   VARIANT      46     46       T -> K (IN STRAIN ML 308-225).
FT   VARIANT     157    157       W -> L (IN STRAIN ML 308-225).
FT   VARIANT     188    188       V -> F (IN STRAIN ML 308-225).
```

```
FT   VARIANT     303    305       ATN -> STI (IN STRAIN ML 308-225).
FT   VARIANT     320    320       A -> T (IN STRAIN ML 308-225).
FT   VARIANT     372    372       N -> Q (IN STRAIN ML 308-225).
FT   VARIANT     493    493       E -> V (IN STRAIN ML 308-225).
FT   VARIANT     497    497       S -> N (IN STRAIN ML 308-225).
FT   VARIANT     585    585       H -> Y (IN STRAIN ML 308-225).
FT   VARIANT     709    709       E -> K (IN STRAIN ML 308-225).
FT   VARIANT     721    721       M -> T (IN STRAIN ML 308-225).
FT   VARIANT     751    753       GHL -> SHF (IN STRAIN ML 308-225).
FT   VARIANT     803    803       S -> P (IN STRAIN ML 308-225).
FT   VARIANT     815    815       A -> V (IN STRAIN ML 308-225).
FT   VARIANT     824    824       C -> S (IN STRAIN ML 308-225).
FT   VARIANT     829    835       LNLVHTS -> MNLIYNA (IN STRAIN ML 308-
FT                                225).
FT   VARIANT     845    847       QGT -> LGA (IN STRAIN ML 308-225).
FT   VARIANT     855    855       S -> T (IN STRAIN ML 308-225).
FT   VARIANT     888    888       Q -> L (IN STRAIN ML 308-225).
FT   VARIANT    1025   1025       S -> I (IN STRAIN ML 308-225).
FT   CONFLICT     61     63       ETV -> TTT (IN REF. 5).
SQ   SEQUENCE   1039 AA;  106841 MW;  5170D647C8DEEBE0 CRC64;

  Query Match             7.8%;  Score 152.5;  DB 1;  Length 1039;
  Best Local Similarity   25.2%;  Pred. No. 0.23;
  Matches  100;  Conservative   50;  Mismatches  126;  Indels  121;  Gaps   27;

Qy          3 AASVKDVLN-AGWNIKGVKPGTTASDNVDFVRT----------YDTVEFLSADTKTTTVN 51
              | :   |:|  || :  ||||| |:| |  | |           || :|:  |   ||:|
Db        162 AIATGTVINDKGWQV--VKPGTVATDTV--VNTGAEGGPDAENGDTGQFVRGDAVRTTIN 217

Qy         52 VESKDNGKKTEVKIG-AKTSVIKEKDGKLVTGKDKGENGSSTDE----GEGLV----TAK 102
                ||::     | | |:|:          | |:  :| : |      |   |    ||
Db        218 ----KNGRQIVRAEGTANTTVV-------YAGGDQTVHGHALDTTLNGGYQYVHNGGTAS 266

Qy        103 EVIDAVNKAGWRM-------KTTTANGQTGQADKFETVTSGTNVTFASG----KGTTATV 151
              : :  ||  ||::          || | | |:  : :   : ||||   |      | |||
Db        267 DTV--VNSDGWQIVKNGGVAGNTTVN-QKGRL-QVDAGGTATNVTLKQGGALVTSTAATV 322

Qy        152 SKDDQG-NITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDET 210
              :    | |    : |  | | ||  |:| |  ||          |::|: : :
Db        323 T----GINRLGAFSVVEGKADNV-VLENGG-RLD----------VLTGHTATN------- 359

Qy        211 VNINAGNNIEITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGDALNVGSKKDNKPV 270
              :: |  ::: |||       |: |     ||:| |     |  |  |   |:: | |
Db        360 TRVDDGGTLDV-RNG-----GTATT-----VSMGNGG---VLLADSGAAVSGTRSDGK-- 403

Qy        271 RITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAYLPGKS 330
              |  :  |    :   || :  ||          ||      |
Db        404 -----AFSIGGGQADALMLEKGSSFTLN----------AGDTATDTT------------ 435

Qy        331 MMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGN 367
              : || :    |   | :::::|    : |    |
Db        436 ---VNGGLFTARGGTLAGTTTLNNGAILTLSGKTVNN 469


RESULT 3
OMPB_RICCN
ID   OMPB_RICCN     STANDARD;      PRT;  1655 AA.
AC   Q9KKA3; Q9KK98; Q9XC45;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Outer membrane protein B precursor (168 kDa surface-layer protein)
DE   (Surface protein antigen) (Cell surface antigen 5) (Sca5) (rOmpB)
DE   (rOmp B) [Contains: 120 kDa surface-exposed protein (Surface protein
DE   antigen) (120 kDa outer membrane protein ompB); 32 kDa beta peptide].
GN   OMPB OR RC1085.
OS   Rickettsia conorii.
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC   Rickettsiaceae; Rickettsieae; Rickettsia.
OX   NCBI_TaxID=781;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Malish 7;
RX   MEDLINE=21442074; PubMed=11557893;
RA   Ogata H., Audic S., Renesto-Audiffren P., Fournier P.-E., Barbe V.,
RA   Samson D., Roux V., Cossart P., Weissenbach J., Claverie J.-M.,
RA   Raoult D.;
RT   "Mechanisms of evolution in Rickettsia conorii and R. prowazekii.";
RL   Science 293:2093-2098(2001).
RN   [2]
RP   SEQUENCE OF 33-1649 FROM N.A.
RC   STRAIN=Indian tick typhus, and Malish 7;
RX   MEDLINE=20393643; PubMed=10939649;
RA   Roux V., Raoult D.;
RT   "Phylogenetic analysis of members of the genus Rickettsia using the
RT   gene coding the outer-membrane protein rOmpB (ompB).";
RL   Int. J. Syst. Evol. Microbiol. 50:1449-1455(2000).
RN   [3]
RP   SEQUENCE OF 353-1655 FROM N.A.
RC   STRAIN=Malish 7;
RA   Stenos J., Walker D.;
RT   "The rickettsial outer membrane protein A and B genes of Rickettsia
RT   australis, the most divergent rickettsia of the spotted fever group.";
RL   Submitted (MAY-1999) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: THE 120 KDA SURFACE-EXPOSED PROTEIN IS A MAJOR
CC       STRUCTURAL PROTEIN WHICH MAY PLAY A ROLE AS A RICKETTSIAL
CC       VIRULENCE FACTOR AND/OR IMMUNOGEN DURING INFECTION (BY
CC       SIMILARITY).
CC   -!- FUNCTION: THE 32 KDA BETA PEPTIDE MAY SERVE AS A MEMBRANE ANCHOR
CC       (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: CELL WALL. THIS BACTERIUM IS COVERED BY A S-
CC       LAYER WITH HEXAGONAL SYMMETRY (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE RICKETTSIAE OMPA/OMPB FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE008659; AAL03623.1; -.
DR   EMBL; AF123721; AAF34124.1; -.
DR   EMBL; AF123726; AAF34129.1; -.
DR   EMBL; AF149110; AAD39533.1; -.
DR   PIR; E97835; E97835.
DR   InterPro; IPR006315; Autotransport.
DR   InterPro; IPR005546; Autotransporter.
DR   Pfam; PF03797; Autotransporter; 1.
DR   TIGRFAMs; TIGR01414; autotrans_barl; 2.
KW   Antigen; S-layer; Cell wall; Complete proteome.
FT   CHAIN         1   1334       120 kDa SURFACE-EXPOSED PROTEIN.
FT   CHAIN      1335   1655       32 kDa BETA PEPTIDE.
FT   VARIANT      61     61       P -> A (IN STRAIN INDIAN TICK TYPHUS).
FT   VARIANT      75     75       G -> S (IN STRAIN INDIAN TICK TYPHUS).
FT   VARIANT      78     78       K -> N (IN STRAIN INDIAN TICK TYPHUS).
FT   VARIANT     251    251       V -> A (IN STRAIN INDIAN TICK TYPHUS).
FT   VARIANT     413    413       N -> D (IN STRAIN INDIAN TICK TYPHUS).
FT   VARIANT     959    959       I -> V (IN STRAIN INDIAN TICK TYPHUS).
FT   VARIANT     988    988       A -> T (IN STRAIN INDIAN TICK TYPHUS).
FT   VARIANT    1139   1139       R -> L (IN STRAIN INDIAN TICK TYPHUS).
FT   CONFLICT    353    354       KD -> GH (IN REF. 3).
FT   CONFLICT    776    776       F -> S (IN REF. 3).
FT   CONFLICT   1159   1159       E -> D (IN REF. 3).
FT   CONFLICT   1177   1177       G -> S (IN REF. 3).
FT   CONFLICT   1492   1492       H -> R (IN REF. 3).
SQ   SEQUENCE   1655 AA;  168342 MW;  E49E19377D5FCE37 CRC64;

  Query Match             7.8%;  Score 152.5;  DB 1;  Length 1655;
  Best Local Similarity   24.4%;  Pred. No. 0.39;
  Matches  103;  Conservative   49;  Mismatches  183;  Indels   87;  Gaps   21;

Qy         10 LNAGWNIKGVKPGTTASDNVDFVRTYD---TVEFLSADTKTTTVN----VESKDNG---- 58
              ||  :      || ||      | |:|   |:   |||      |    :|:   |
Db        417 LNGNFTGDASNPGNTAG-----VITFDANGTLASASADANVAVTNNITAIEASGAGVVQL 471
```

```
Qy         59 ---KKTEVKIGAKTSVIKEKDGKLVTGKDKG--------ENGSSTDEGEGLVTAKEVIDA 107
                  |:::|   || |  || :: ||             |: | :|   :|     |
Db        472 SGTHAAELRLGNAGSVFKLADGTVINGKVNQTALVGGALAAGTITLDGSATITG----DI 527

Qy        108 VNKAGWRMKTTTANGQTGQADKFETVT---------SGTNVTFASGKGTTATVSKDDQGN 158
                |  |          | |   |  :|:|         :|  : | :  ||    |   | |
Db        528 GNAGG----AAALQGITLANDATKTLTLGGANIIGANGGTINFQANGGTIKLTS--TQNN 581

Qy        159 ITVMYDV-------NVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKM-DET 210
                | | :|:        | || ::   |    |     |  ::  :   |:  ||  : |
Db        582 IVVDFDLAIATDQTGVVDASSLTNAQTLTINGKIGTVGANNKTLGQFNIGSSKTVLSDGD 641

Qy        211 VNIN---AGNN--IEITRNGKNIDIATSMTPQ----FSSV-----SLGAGADAPTLSVDG 256
                | ||    |||  ::   |   |   |:   |    |: |     :|  | :  : :
Db        642 VAINELVIGNNGAVQFAHNTYLITRTTNAAGQGKIIFNPVVNNNTTLATGTNLGSATNPL 701

Qy        257 DALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDG-NARAGIAQA 315
                :| |||       : ||  ||     ||:        |: : | |  | |  :|
Db        702 AEINFGSKGAANVDTVLNVGKGVNL-YATNITTTDA---NVGSFIFNAGGTNIVSG---- 753

Qy        316 IATAGLVQ-------AYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIKGTASGNS 368
                | |  |       |   | ::  :|  |: |    |   |::  |||:     || :
Db        754 --TVGGQQGNKFNTVALDNGTTVKFLGNATFNGNTTIAAN-STLQIGGNYTADFVASADG 810

Qy        369 RG 370
                |
Db        811 TG 812
```

```
RESULT 4
YDBA_ECOLI
ID   YDBA_ECOLI     STANDARD;      PRT;  2003 AA.
AC   P33666; P76087; P76088; P76856; P76857; P76859;
DT   01-FEB-1994 (Rel. 28, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Hypothetical protein ydbA.
GN   YDBA OR B1401/B1405.
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12 / MG1655;
RX   MEDLINE=97426617; PubMed=9278503;
RA   Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA   Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA   Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA   Mau B., Shao Y.;
RT   "The complete genome sequence of Escherichia coli K-12.";
RL   Science 277:1453-1474(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12;
RX   MEDLINE=97251357; PubMed=9097039;
RA   Aiba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K., Itoh T.,
RA   Kasai H., Kashimoto K., Kimura S., Kitakawa M., Kitagawa M.,
RA   Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,
RA   Nakade S., Nakamura Y., Nashimoto H., Nishio Y., Oshima T., Saito N.,
RA   Sampei G., Seki Y., Sivasundaram S., Tagami H., Takeda J.,
RA   Takemoto K., Takeuchi Y., Wada C., Yamamoto Y., Horiuchi T.;
RT   "A 570-kb DNA sequence of the Escherichia coli K-12 genome
RT   corresponding to the 28.0-40.1 min region on the linkage map.";
RL   DNA Res. 3:363-377(1996).
RN   [3]
RP   SEQUENCE OF 464-2003 FROM N.A.
RC   STRAIN=K12;
RX   MEDLINE=92190338; PubMed=1665988;
RA   Moszer I., Glaser P., Danchin A.;
RT   "Multiple IS insertion sequences near the replication terminus in
RT   Escherichia coli K-12.";
RL   Biochimie 73:1361-1374(1991).
CC   -!- SIMILARITY: TO S.TYPHIMURIUM ORF NEAR CYSG (AC P25928).
CC   -!- CAUTION: THIS IS A CONCEPTUAL TRANSLATION; THE GENE CODING FOR
CC       THIS PROTEIN IS INTERRUPTED BY A HYBRID IS2D/IS30C ELEMENT
CC       BETWEEN AMINO ACIDS 839 AND 840.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE000237; AAC74483.1; ALT_SEQ.
DR   EMBL; AE000237; AAC74487.1; ALT_SEQ.
DR   EMBL; D90778; BAA15009.1; ALT_SEQ.
DR   EMBL; D90778; BAA18880.1; ALT_SEQ.
DR   EMBL; D90779; BAA18881.1; ALT_SEQ.
DR   EMBL; X62680; -; NOT_ANNOTATED_CDS.
DR   EcoGene; EG11307; ydbA.
KW   Hypothetical protein; Complete proteome.
FT   CONFLICT    489    489       I -> V (IN REF. 2).
FT   CONFLICT    495    495       I -> V (IN REF. 2).
SQ   SEQUENCE   2003 AA;  205949 MW;  B83A12C8B53220EE CRC64;

  Query Match             7.8%;  Score 152;  DB 1;  Length 2003;
  Best Local Similarity   23.5%;  Pred. No. 0.52;
  Matches  104;  Conservative   55;  Mismatches  157;  Indels  126;  Gaps   24;

Qy         22 GTTASDN---VDFVRTYDTVEFLSADTKT-----------TTVNVESKDNGKKTEVKIG 66
                |  |: |   :|       :| :|||  |           |  |    :||| |   :
Db        179 GRKATINLWQIDEANNTVALEGVSADGATKWQYNHNGELVITGDNATVNNNGKTT---VD 235

Qy         67 AKTSV---IKEKDGKLVTGKDKGENG----------SSTDEGEGLVTAKE------VIDA 107
               | |    |   :||::   |   :|          |:| : :| :|  :       ||
Db        236 GKDSTGTEINGNNGKVIQDGDLDVSGGGHGIDITGDSATVDNKGTMTVTDPESMGIQIDG 295

Qy        108 ----VNKAGWRMKTTTANGQTGQADKFETVT-------------SGTNVTFASGK---- 145
                   ||  |   ::|  || ||      :  |             :|| :   :||
Db        296 DKAIVNNEG---ESTITNGGTGTQINGDDATANNNGKTTVDGKDSTGTEINGNNGKVIQD 352

Qy        146 ---------------GTTATVSKDDQGNITVMYDVNVGDALNVNQ--LQNSGWNLDSKAV 188
                            | :|||  |::| :||    ::|  :: :|  : :| |   :|
Db        353 GDLDVSGGGHGIDITGDSATV--DNKGTMTVTDPESIGIQVDGDQAVVNNEG---ESAIT 407

Qy        189 AGSSGKVISGN--VSPSKGK-----MDETVNINAGNN--------IEITRNGKNIDI-AT 232
                | :|  |:|:   : : ||      | |    ||||        :::: |  |||
Db        408 NGGTGTQINGDDATANNNGKTTVDGKDSTGTEIAGNNGKVIQDGDLDVSGGGHGIDITGD 467

Qy        233 SMTPQFSSVSLGAGADAPTLSVDGDALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKG 292
                | |              ::  : :|||   | :: ::    |||   |          |: |
Db        468 SATVDNKGTMTVTDPESIGIQIDGDQAIVNNEGES---TITNGGTG---------TQING 515

Qy        293 VAQNLNNR-IDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSS 351
                  ||     |||    |  :      |:|   | | |:   ||          | |  :
Db        516 NDATANNSGKTTVDGKDSTG-TKIAGNIGIVN--LDG-SLTVTGG---------AHGVEN 562

Qy        352 ISDGGNWIIKGTASGNSRGHFG 373
                | | |    ||    :  |  |
Db        563 IGDNGTVNNKGDIVVSDTGSIG 584
```

```
RESULT 5
120K_RICRI
ID   120K_RICRI     STANDARD;      PRT;  1300 AA.
AC   P14914;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   120 kDa surface-exposed protein.
GN   P120.
OS   Rickettsia rickettsii.
```

```
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC   Rickettsiaceae; Rickettsieae; Rickettsia.
OX   NCBI_TaxID=783;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=R;
RX   MEDLINE=90136087; PubMed=2515418;
RA   Gilmore R.D. Jr., Joste N., McDonald G.A.;
RT   "Cloning, expression and sequence analysis of the gene encoding the
RT   120 kD surface-exposed protein of Rickettsia rickettsii.";
RL   Mol. Microbiol. 3:1579-1586(1989).
CC   -!- FUNCTION: MAJOR STRUCTURAL PROTEIN WHICH MAY PLAY A ROLE AS
CC       RICKETTSIAL VIRULENCE FACTOR AND/OR IMMUNOGEN DURING INFECTION.
CC   -!- SUBCELLULAR LOCATION: CELL WALL. THIS BACTERIUM IS COVERED BY A
CC       S-LAYER WITH HEXAGONAL SYMMETRY.
CC   -!- MISCELLANEOUS: ITS C-TERMINUS POTENTIALLY MAY BEAR THE EPITOPES
CC       CONFERRING ANTIGENICITY TO THE PROTEIN.
CC   -!- SIMILARITY: BELONGS TO THE RICKETTSIAE OMPA/OMPB FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X16353; CAA34402.1; -.
DR   PIR; S07575; S07575.
DR   InterPro; IPR006315; Autotransport.
DR   InterPro; IPR005546; Autotransporter.
DR   Pfam; PF03797; Autotransporter; 1.
DR   TIGRFAMs; TIGR01414; autotrans_barl; 2.
KW   Antigen; Glycoprotein; Cell wall; S-layer.
FT   CARBOHYD      7      7       N-LINKED (POTENTIAL).
FT   CARBOHYD     66     66       N-LINKED (POTENTIAL).
FT   CARBOHYD     86     86       N-LINKED (POTENTIAL).
FT   CARBOHYD    103    103       N-LINKED (POTENTIAL).
FT   CARBOHYD    147    147       N-LINKED (POTENTIAL).
FT   CARBOHYD    268    268       N-LINKED (POTENTIAL).
FT   CARBOHYD    330    330       N-LINKED (POTENTIAL).
FT   CARBOHYD    375    375       N-LINKED (POTENTIAL).
FT   CARBOHYD    415    415       N-LINKED (POTENTIAL).
FT   CARBOHYD    424    424       N-LINKED (POTENTIAL).
FT   CARBOHYD    430    430       N-LINKED (POTENTIAL).
FT   CARBOHYD    436    436       N-LINKED (POTENTIAL).
FT   CARBOHYD    444    444       N-LINKED (POTENTIAL).
FT   CARBOHYD    515    515       N-LINKED (POTENTIAL).
FT   CARBOHYD    547    547       N-LINKED (POTENTIAL).
FT   CARBOHYD    593    593       N-LINKED (POTENTIAL).
FT   CARBOHYD    655    655       N-LINKED (POTENTIAL).
FT   CARBOHYD    698    698       N-LINKED (POTENTIAL).
FT   CARBOHYD    710    710       N-LINKED (POTENTIAL).
FT   CARBOHYD    799    799       N-LINKED (POTENTIAL).
FT   CARBOHYD    800    800       N-LINKED (POTENTIAL).
FT   CARBOHYD    826    826       N-LINKED (POTENTIAL).
FT   CARBOHYD    844    844       N-LINKED (POTENTIAL).
FT   CARBOHYD    861    861       N-LINKED (POTENTIAL).
FT   CARBOHYD    879    879       N-LINKED (POTENTIAL).
FT   CARBOHYD    920    920       N-LINKED (POTENTIAL).
FT   CARBOHYD    926    926       N-LINKED (POTENTIAL).
FT   CARBOHYD   1116   1116       N-LINKED (POTENTIAL).
FT   CARBOHYD   1128   1128       N-LINKED (POTENTIAL).
FT   CARBOHYD   1140   1140       N-LINKED (POTENTIAL).
FT   CARBOHYD   1146   1146       N-LINKED (POTENTIAL).
FT   CARBOHYD   1211   1211       N-LINKED (POTENTIAL).
SQ   SEQUENCE   1300 AA;  132801 MW;  E09E52C3F647243D CRC64;

  Query Match             7.7%;  Score 150;  DB 1;  Length 1300;
  Best Local Similarity  23.0%;  Pred. No. 0.41;
  Matches  104;  Conservative   50;  Mismatches  172;  Indels  126;  Gaps   22;

Qy          1 NRAASVK--DVLNAGWNIKG--VKPGTTASDNVDFVRTYD---TVEFLSADTKTTTVN-- 51
              | || :|  : :    |  |    || ||      | |:|   |:|  |||      |
Db         50 NLAAQIKVPNAITLTGNFTGDASNPGNTAG-----VITFDANGTLESASADANVAVTNNI 104

Qy         52 --VESKDNG-------KKTEVKIGAKTSVIKEKDGKLVTGKDK------GENGSSTDEGE 96
                :|:   |          |:::|   |: |  || :: ||        |   : |   :
Db        105 TAIEASGAGVVQLSGTHAAELRLGNAGSIFKLADGTVINGKVNQTALVGGALAAGTITLD 164

Qy         97 GLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVT-SGTNVTFASG-------KGTT 148
              |  |    |     |    : | ||      |  :|:|  | |:  | |        | |
Db        165 GSATITGDIGNAGGAAALQRITLAN------DAKKTLTLGGANIIGAGGGTIDLQANGGT 218

Qy        149 ATVSKDDQGNITVMYDV-------NVGDALNVNQLQNSGWN-------LDSKAVA----G 190
                ::   | || | :|:        | || ::   |     |        ::| :     |
Db        219 IKLT-STQNNIVVDFDLAIATDQTGVVDASSLTNAQTLTINGKIGTIGANNKTLGQFNIG 277

Qy        191 SSGKVIS-GNV--------------------------SPSKGKMDETVNINAGNNIEIT 222
              ||  |:| |||                        :  :||:     :| |  :
Db        278 SSKTVLSNGNVAINELVIGNDGAVQFAHDTYLITRTTNAAGQGKIIFNPVVNNGTTLAA- 336

Qy        223 RNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGDALNVGSKKDNKPVRIT----NVAPG 278
                | |:  ||:   : :  | |   |          ||||   :     ||    ||
Db        337 --GTNLGSATNPLAEINFGSKGVNVDT--------VLNVGEGVNLYATNITTTDANVGSF 386

Qy        279 VKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGT 338
              |     ||:     |     |: : |                  |   | ::  :|  |
Db        387 VFNAGGTNIVS-GTVGGQQGNKFNTV------------------ALENGTTVKFLGNAT 426

Qy        339 YRGEAGYAIGYSSISDGGNWIIKGTASGNSRG 370
              : |    |   |::  |||:     || :  |
Db        427 FNGNTTIAAN-STLQIGGNYTADCVASADGTG 457


RESULT 6
OMPB_RICRI
ID   OMPB_RICRI     STANDARD;      PRT;  1654 AA.
AC   Q53047;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Outer membrane protein B precursor (168 kDa surface-layer protein)
DE   (Surface protein antigen) (Cell surface antigen 5) (Sca5) (rOmpB)
DE   (rOmp B) [Contains: 120 kDa surface-exposed protein (Surface protein
DE   antigen) (120 kDa outer membrane protein ompB); 32 kDa beta peptide].
GN   OMPB.
OS   Rickettsia rickettsii.
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC   Rickettsiaceae; Rickettsieae; Rickettsia.
OX   NCBI_TaxID=783;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=R;
RX   MEDLINE=92167802; PubMed=1724278;
RA   Gilmore R.D. Jr., Cieplak W. Jr., Policastro P.F., Hackstadt T.;
RT   "The 120 kilodalton outer membrane protein (rOmp B) of Rickettsia
RT   rickettsii is encoded by an unusually long open reading frame:
RT   evidence for protein processing from a large precursor.";
RL   Mol. Microbiol. 5:2361-2370(1991).
RN   [2]
RP   SEQUENCE OF 279-1654 FROM N.A.
RC   STRAIN=R;
RX   MEDLINE=90136087; PubMed=2515418;
RA   Gilmore R.D. Jr., Joste N., McDonald G.A.;
RT   "Cloning, expression and sequence analysis of the gene encoding the
RT   120 kD surface-exposed protein of Rickettsia rickettsii.";
RL   Mol. Microbiol. 3:1579-1586(1989).
CC   -!- FUNCTION: THE 120 kDa SURFACE-EXPOSED PROTEIN IS A MAJOR
CC       STRUCTURAL PROTEIN WHICH MAY PLAY A ROLE AS A RICKETTSIAL
CC       VIRULENCE FACTOR AND/OR IMMUNOGEN DURING INFECTION.
CC   -!- SUBCELLULAR LOCATION: CELL WALL. THIS BACTERIUM IS COVERED BY A S-
CC       LAYER WITH HEXAGONAL SYMMETRY.
CC   -!- SIMILARITY: BELONGS TO THE RICKETTSIAE OMPA/OMPB FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
```

```
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; X16353; CAA34403.1; -.
DR   PIR; S18227; S18227.
DR   InterPro; IPR006315; Autotransport.
DR   InterPro; IPR005546; Autotransporter.
DR   Pfam; PF03797; Autotransporter; 1.
DR   TIGRFAMs; TIGR01414; autotrans_barl; 2.
KW   Antigen; S-layer; Cell wall.
FT   CHAIN         1   1333       120 kDa SURFACE-EXPOSED PROTEIN.
FT   CHAIN      1334   1654       32 kDa BETA PEPTIDE.
FT   DOMAIN     1181   1188       POLY-THR.
SQ   SEQUENCE   1654 AA;  168184 MW;  D7AB70FB7087F618 CRC64;

  Query Match             7.7%;  Score 150;  DB 1;  Length 1654;
  Best Local Similarity   23.0%;  Pred. No. 0.54;
  Matches  104;  Conservative   50;  Mismatches  172;  Indels  126;  Gaps   22;

Qy          1 NRAASVK--DVLNAGWNIKG--VKPGTTASDNVDFVRTYD---TVEFLSADTKTTTVN--  51
              | || :|  : :    | |    || ||      | |:|   |:|  |||      |
Db        404 NLAAQIKVPNAITLTGNFTGDASNPGNTAG-----VITFDANGTLESASADANVAVTNNI 458

Qy         52 --VESKDNG-------KKTEVKIGAKTSVIKEKDGKLVTGKDK------GENGSSTDEGE 96
                :|:   |       |:::|   |: |  || :: ||       |   : |   :
Db        459 TAIEASGAGVVQLSGTHAAELRLGNAGSIFKLADGTVINGKVNQTALVGGALAAGTITLD 518

Qy         97 GLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVT-SGTNVTFASG-------KGTT 148
              |  |    |     |    : | ||       |  :|:|  | |:  | |        | |
Db        519 GSATITGDIGNAGGAAALQRITLAN------DAKKTLTLGGANIIGAGGGTIDLQANGGT 572

Qy        149 ATVSKDDQGNITVMYDV-------NVGDALNVNQLQNSGWN-------LDSKAVA----G 190
               ::   | || | :|:        | || ::   |    |        ::| :     |
Db        573 IKLT-STQNNIVVDFDLAIATDQTGVVDASSLTNAQTLTINGKIGTIGANNKTLGQFNIG 631

Qy        191 SSGKVIS-GNV--------------------------SPSKGKMDETVNINAGNNIEIT 222
              ||  |:| |||                        :  :||:     :| |  :
Db        632 SSKTVLSNGNVAINELVIGNDGAVQFAHDTYLITRTTNAAGQGKIIFNPVVNNGTTLAA- 690

Qy        223 RNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGDALNVGSKKDNKPVRIT----NVAPG 278
               | |:  ||:   : :  | |   |          ||||   :     ||    ||
Db        691 --GTNLGSATNPLAEINFGSKGVNVDT--------VLNVGEGVNLYATNITTTDANVGSF 740

Qy        279 VKEGDVTNVAQLKGVAQNLNNRIDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGT 338
              |     ||:     |     |: : |                    |   | ::  :|  |
Db        741 VFNAGGTNIVS-GTVGGQQGNKFNTV------------------ALENGTTVKFLGNAT 780

Qy        339 YRGEAGYAIGYSSISDGGNWIIKGTASGNSRG 370
              : |    |   |::  |||:     || :  |
Db        781 FNGNTTIAAN-STLQIGGNYTADCVASADGTG 811


RESULT 7
FLIC_SHIFL
ID   FLIC_SHIFL     STANDARD;      PRT;   550 AA.
AC   Q08860;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Flagellin.
GN   FLIC OR SF1966.
OS   Shigella flexneri.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Shigella.
OX   NCBI_TaxID=623;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=IID642;
RX   MEDLINE=94335647; PubMed=8057852;
RA   Tominaga A., Mahmoud M.A.-H., Mukaihara T., Enomoto M.;
RT   "Molecular characterization of intact, but cryptic, flagellin genes
RT   in the genus Shigella.";
RL   Mol. Microbiol. 12:277-285(1994).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=301 / Serotype 2a;
RX   MEDLINE=22272406; PubMed=12384590;
RA   Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
RA   Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
RA   Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
RA   Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA   Yu J.;
RT   "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
RT   through comparison with genomes of Escherichia coli K12 and O157.";
RL   Nucleic Acids Res. 30:4432-4441(2002).
CC   -!- FUNCTION: FLAGELLIN IS THE SUBUNIT PROTEIN WHICH POLYMERIZES TO
CC       FORM THE FILAMENTS OF BACTERIAL FLAGELLA.
CC   -!- SIMILARITY: BELONGS TO THE BACTERIAL FLAGELLIN FAMILY.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; D16819; BAA04093.1; -.
DR   EMBL; AE015215; AAN43516.1; -.
DR   PIR; S44980; S44980.
DR   InterPro; IPR001029; Flagellin_C.
DR   InterPro; IPR001492; FlagellinN.
DR   Pfam; PF00700; Flagellin_C; 1.
DR   Pfam; PF00669; Flagellin_N; 1.
DR   PRINTS; PR00207; FLAGELLIN.
KW   Flagella.
SQ   SEQUENCE   550 AA;  56636 MW;  CC921C9A8EF200B6 CRC64;

  Query Match             7.4%;  Score 144;  DB 1;  Length 550;
  Best Local Similarity   21.5%;  Pred. No. 0.33;
  Matches   76;  Conservative   55;  Mismatches  167;  Indels   56;  Gaps   14;

Qy         22 GTTASDNVDFVR-----TYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKD 76
              || :  ::| ::       | :: :|  |:   ||| :||   |  :::||        |
Db        102 GTNSDSDLDSIQDEIKSRLDEIDRVSGQTQFNGVNVLAKDGSMK--IQVGANDGQTITID 159

Qy         77 GKLVTGKDKGENGSSTDEGEGL---VTAKEVIDAVNKAGWRMKTTTANGQTG--QADKFE 131
              | :     | || : : |  :     :|  : | |      | | : |      :|
Db        160 LKKIDSDTLGLNGFNVNGGGAVANTAASKADLVAANATVVGNKYTVSAGYDAAKASDLLA 219

Qy        132 TVTSGTNV--TFASGKGTTATVS--KDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKA 187
              |: |  |  |  :| || |: :  | |  : :  :|     | :| :    |
Db        220 GVSDGDTVQATINNGFGTAASATNYKYDSASKSYSFDTTTASAADVQKYLTPG------- 272

Qy        188 VAGSSGK---VISGNVSPSKGKMDETVNINAGNNI------EITRNGKNIDIATSMTPQF 238
                | : |    | |:    :   |  :  : |: :       :|:||       |:|
Db        273 -VGDTAKGTITIDGSAQDVQISSDGKITASNGDKLYIDTTGRLTKNGS----GASLTEAS 327

Qy        239 SSVSLGAGADAPTLSVDGDALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLN 298
              |         | |: : | :::  :     |  ||    | |   |    | | |: : |
Db        328 LSTLAANNTKATTIDIGGTSISF-TGNSTTPDTITYSVTGAK---VDQAAFDKAVSTSGN 383

Qy        299 N--------RIDNVDGNARAGI--------AQAIATAGLVQAYLPGKSMMAIGGG 337
              |         ::   |    |:        :|: |:  |  ||     :  | |
Db        384 NVDFTTAGYSVNGTTGAVTKGVDSVYVDNNEALTTSDTVDFYLQDDGSVTNGSG 437


RESULT 8
BIGA_SALTY
ID   BIGA_SALTY     STANDARD;      PRT;  1953 AA.
AC   P25927; P25928; Q9XCQ3;
DT   01-MAY-1992 (Rel. 22, Created)
```

```
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Putative surface-exposed virulence protein bigA precursor.
GN   BIGA OR STM3478.
OS   Salmonella typhimurium.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Salmonella.
OX   NCBI_TaxID=602;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 14028;
RA   Stojiljkovic I., Valentine P., Heffron F.;
RT   "Salmonella typhimurium rhs homolog.";
RL   Submitted (MAR-1999) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX   MEDLINE=21534948; PubMed=11677609;
RA   McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA   Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA   Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA   Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA   Waterston R., Wilson R.K.;
RT   "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT   LT2.";
RL   Nature 413:852-856(2001).
RN   [3]
RP   SEQUENCE OF 1-765 FROM N.A.
RC   STRAIN=LT2;
RX   MEDLINE=91100301; PubMed=1987123;
RA   Wu J.Y., Siegel L.M., Kredich N.M.;
RT   "High-level expression of Escherichia coli NADPH-sulfite reductase:
RT   requirement for a cloned cysG plasmid to overcome limiting siroheme
RT   cofactor.";
RL   J. Bacteriol. 173:325-333(1991).
CC   -!- CAUTION: Ref.3 sequence differs from that shown due to frameshifts
CC       in positions 414 and 732.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF133696; AAD39458.1; -.
DR   EMBL; AE008859; AAL22340.1; -.
DR   EMBL; M64606; AAA27042.1; ALT_FRAME.
DR   EMBL; M64606; AAA27043.1; ALT_FRAME.
DR   StyGene; SG10437; bigA.
KW   Virulence; Repeat; Signal; Complete proteome.
FT   SIGNAL        1     27       POTENTIAL.
FT   CHAIN        28   1953       PUTATIVE SURFACE-EXPOSED VIRULENCE
FT                                PROTEIN BIGA.
FT   DOMAIN      101    252       15 X 11 AA TANDEM REPEATS.
FT   REPEAT      101    103       1 (INCOMPLETE).
FT   REPEAT      104    113       2 (INCOMPLETE).
FT   REPEAT      114    122       3 (INCOMPLETE).
FT   REPEAT      123    133       4.
FT   REPEAT      134    144       5.
FT   REPEAT      145    155       6.
FT   REPEAT      156    166       7.
FT   REPEAT      167    177       8.
FT   REPEAT      178    188       9.
FT   REPEAT      189    199       10.
FT   REPEAT      200    210       11.
FT   REPEAT      211    221       12.
FT   REPEAT      222    232       13.
FT   REPEAT      233    243       14.
FT   REPEAT      244    252       15 (INCOMPLETE).
FT   CONFLICT    207    207       D -> DRGDDDVTPPDD (IN REF. 1).
FT   CONFLICT    514    514       A -> R (IN REF. 3).
FT   CONFLICT   1698   1698       D -> N (IN REF. 1).
FT   CONFLICT   1795   1798       QYLE -> ITLQ (IN REF. 1).
FT   CONFLICT   1836   1837       SA -> T (IN REF. 1).
SQ   SEQUENCE   1953 AA;  200150 MW;  611B3F1C954D91AE CRC64;

  Query Match             7.3%;  Score 142;  DB 1;  Length 1953;
  Best Local Similarity   23.3%;  Pred. No. 1.8;
  Matches   91;  Conservative   48;  Mismatches  144;  Indels  108;  Gaps   20;

Qy         13 GWNI----KGVKPGTTASDNVDFVRTYDTVEFL-SADTKTTTVNVESKDNGKKTEVKIGA 67
              ||::      |  ||  :        ||:  :|  ||  || :: : : :     : |
Db        310 GWDVTDANAAVIEGTQEN---GLYWKYDSRGYLIIADDNTTVISGDDQAHNSDRGMDISG 366

Qy         68 KTSVIKEKDGKLVTGKDKGEN-----GSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANG 122
                 ::: | :::| |:  |       | ||   |:|                    : :|
Db        367 -----QDRTGVIISG-DRTVNTLTGDSSVTDGATGMV-----------------ISGDG 402

Qy        123 QTGQADKFETVTSGTNVTFASGKGTTATVSKD---DQGNITVMYDVNVGDALNVNQLQNS 179
               |       || : |     || |||   : |     |   :: | :     |      |
Db        403 TTNTISGHSTVDNATG-ALISGNGTTTNFAGDIAVSGGGTAIIIDGDNATIKNTGTSDIS 461

Qy        180 GWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFS 239
               |       |||:| || || :      | |:  : |    || :   || | | |
Db        462 G--------AGSTGTVIDGNNARVNNDGDMTIT-DGGTGGHITGDNVVIDNAGSTTV--- 509

Qy        240 SVSLGAGADAPTLSVDGD---ALNVGSKKDNKPVRITNVAPGVK-EGDVTNVAQLKGVAQ 295
                 :|||| | ::||    :| |::       |:  | | : :|| :      :|
Db        510 -----SGADATALYIEGDNALVINEGNQ------TISGGAVGTRIDGDDAHTTNTGDIA- 557

Qy        296 NLNNRIDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDG 355
                    |||   | :      | ||                        |   ::||
Db        558 --------VDGAGSAAVIINGDNGSLTQA----------------------GDLLVTDG 586

Qy        356 GNWIIK-GTASGNSRGHFGASA-----SVGY 380
                  ||  ||  ||   : | :       |||:
Db        587 AMGIITYGT--GNEAKNTGNATVRDADSVGF 615


RESULT 9
ALYS_ENTFA
ID   ALYS_ENTFA     STANDARD;      PRT;   737 AA.
AC   P37710;
DT   01-OCT-1994 (Rel. 30, Created)
DT   15-SEP-2003 (Rel. 42, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Autolysin precursor (EC 3.2.1.-) (Peptidoglycan hydrolase)
DE   (Beta-glycosidase).
GN   EF0799.
OS   Enterococcus faecalis (Streptococcus faecalis).
OC   Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.
OX   NCBI_TaxID=1351;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91358349; PubMed=1679432;
RA   Beliveau C., Potvin C., Trudel J., Asselin A., Bellemare G.;
RT   "Cloning, sequencing, and expression in Escherichia coli of a
RT   Streptococcus faecalis autolysin.";
RL   J. Bacteriol. 173:5619-5623(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=V583 / ATCC 700802;
RX   MEDLINE=22550857; PubMed=12663927;
RA   Paulsen I.T., Banerjei L., Myers G.S.A., Nelson K.E., Seshadri R.,
RA   Read T.D., Fouts D.E., Eisen J.A., Gill S.R., Heidelberg J.F.,
RA   Tettelin H., Dodson R.J., Umayam L., Brinkac L., Beanan M.,
RA   Daugherty S., DeBoy R.T., Durkin S., Kolonay J., Madupu R., Nelson W.,
RA   Vamathevan J., Tran B., Upton J., Hansen T., Shetty J., Khouri H.,
RA   Utterback T., Radune D., Ketchum K.A., Dougherty B.A., Fraser C.M.;
RT   "Role of mobile DNA in the evolution of vancomycin-resistant
RT   Enterococcus faecalis.";
RL   Science 299:2071-2074(2003).
CC   -!- FUNCTION: Hydrolyzes the cell wall of E.faecalis and
CC       M.lysodeikticus. May play an important role in cell wall growth
```

```
CC       and cell separation.
CC   -!- SUBCELLULAR LOCATION: Secreted (Probable).
CC   -!- DOMAIN: LYSM REPEATS ARE THOUGHT TO BE INVOLVED IN PEPTIDOGLYCAN
CC       BINDING.
CC   -!- SIMILARITY: BELONGS TO FAMILY 73 OF GLYCOSYL HYDROLASES.
CC   -!- SIMILARITY: Contains 6 LysM repeats.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M58002; AAA67325.1; -.
DR   EMBL; AE016949; AAO80613.1; -.
DR   PIR; A38109; A38109.
DR   TIGR; EF0799; -.
DR   InterPro; IPR002901; Amidase_4.
DR   InterPro; IPR002482; LysM.
DR   Pfam; PF01832; Amidase_4; 1.
DR   Pfam; PF01476; LysM; 5.
DR   SMART; SM00257; LysM; 6.
DR   SMART; SM00047; LYZ2; 1.
KW   Hydrolase; Glycosidase; Bacteriolytic enzyme; Cell wall;
KW   Cell division; Septation; Repeat; Signal.
FT   SIGNAL       1     53       POTENTIAL.
FT   CHAIN       54    737       AUTOLYSIN.
FT   REPEAT     363    405       LYSM 1.
FT   REPEAT     431    473       LYSM 2.
FT   REPEAT     499    541       LYSM 3.
FT   REPEAT     567    609       LYSM 4.
FT   REPEAT     633    675       LYSM 5.
FT   REPEAT     695    737       LYSM 6.
FT   CONFLICT    85     85       T -> I (IN REF. 1).
FT   CONFLICT   118    118       A -> V (IN REF. 1).
FT   CONFLICT   143    143       A -> T (IN REF. 1).
FT   CONFLICT   417    417       S -> N (IN REF. 1).
FT   CONFLICT   449    449       S -> T (IN REF. 1).
FT   CONFLICT   476    476       A -> T (IN REF. 1).
FT   CONFLICT   484    484       N -> S (IN REF. 1).
FT   CONFLICT   567    632       MISSING (IN REF. 1).
SQ   SEQUENCE   737 AA;  77025 MW;  ABB16BD506AC7507 CRC64;

  Query Match             7.2%;  Score 139.5;  DB 1;  Length 737;
  Best Local Similarity   23.5%;  Pred. No. 0.81;
  Matches   77;  Conservative   40;  Mismatches  145;  Indels   65;  Gaps   15;

Qy        8 DVLNAGWNIKG--VKPGTTASDNVDFVRTY------DTVEFLSADTKTTTVNVESKDNGK 59
            |  ::| |  |  | |||  |:|     ||      ||:  ::|    :  |: | :
Db      333 DTPSSGGNTGGGTVNPGTGGSNNQSGTNTYYTVKSGDTLNKIAAQYGVSVANLRSWNGIS 392

Qy       60 KTEVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDE-GEGLVTAKEVIDAVNKAGWRMKTT 118
              : :| |  | |   |   ||   |  ||: :: |       :  | :||   :   :
Db      393 GDLIFVGQKLIVKKGASGN--TG-GSGSGGSNNNQSGTNTYYTVKSGDTLNKIAAQYGVS 449

Qy      119 TANGQTGQADKFETVTSGTNVTF---ASGK--GTTATVSKDDQGNITVMYDVNVGDALN- 172
             || ::      : :  |  :     |||   |:    | ::|      | :  || ||
Db      450 VANLRSWNGISGDLIFVGQKLIVKKGASGNTGGSNNGGSNNNQSGTNTYYTIKSGDTLNK 509

Qy      173 ------VNQLQNSGWNLDSKAVAGSSGKVI-----SGNVSPSKGKMDETVNINAGNNIEI 221
                 |:      ||  |  :  :  |:|     |||   |           | |:|
Db      510 IAAQYGVSVANLRSWNGISGDLIFAGQKIIVKKGTSGNTGGSS---------NGGSN--N 558

Qy      222 TRNGKNIDIATSMTPQFSSVSLGAGADAPTL----SVDGDAL-----------NVGSKK 265
             ::| |           : :|   |     |    :: | :           | ||
Db      559 NQSGTNTYYTIKSGDTLNKISAQFGVSVANLQAWNNISGSLIFAGQKIIVKKGANSGSTN 618

Qy      266 DNKPVRITNVAPG------VKEGDVTN 286
              |||   ||   |      :| ||  |
Db      619 TNKP---TNNGGGATTSYTIKSGDTLN 642
```

```
RESULT 10
HLYA_SERMA
ID   HLYA_SERMA     STANDARD;      PRT;  1608 AA.
AC   P15320;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   01-NOV-1990 (Rel. 16, Last annotation update)
DE   Hemolysin precursor.
GN   SHLA.
OS   Serratia marcescens.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Serratia.
OX   NCBI_TaxID=615;
RN   [1]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 31-40.
RC   STRAIN=SN8;
RX   MEDLINE=88257037; PubMed=3290200;
RA   Poole K., Schiebel E., Braun V.;
RT   "Molecular characterization of the hemolysin determinant of Serratia
RT   marcescens.";
RL   J. Bacteriol. 170:3177-3188(1988).
CC   -!- FUNCTION: BACTERIAL HEMOLYSINS ARE EXOTOXINS THAT ATTACK BLOOD
CC       CELL MEMBRANES AND CAUSE CELL RUPTURE BY MECHANISMS NOT CLEARLY
CC       DEFINED.
CC   -!- FUNCTION: CELL-BOUND HEMOLYSIN, WHICH RELEASES HEME-IRON FROM
CC       ERYTHROCYTES BY INTERACTION WITH THE ERYTHROCYTE MEMBRANE. SHLA
CC       REQUIRES SHLB FUNCTION.
CC   -!- SUBCELLULAR LOCATION: Outer membrane.
CC   -!- SIMILARITY: TO P.MIRABILIS HEMOLYSIN (HPMA).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M22618; AAA50323.1; -.
DR   PIR; A28182; A28182.
KW   Hemolysis; Toxin; Outer membrane; Signal.
FT   SIGNAL       1     30
FT   CHAIN       31   1608       HEMOLYSIN.
SQ   SEQUENCE   1608 AA;  165078 MW;  D669B476FE7DAD51 CRC64;

  Query Match             7.1%;  Score 139;  DB 1;  Length 1608;
  Best Local Similarity   21.6%;  Pred. No. 2.1;
  Matches  109;  Conservative   52;  Mismatches  176;  Indels  168;  Gaps   21;

Qy        1 NRAASVKDVLNA----GWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKD 56
            |  ::  | ::|     :||      |::|   |      |   | :||  | |: :  |
Db      603 NALSTTVDKIDARTGTAFNI------TSSSHKADNSYQSSTASELKSDTNLTLVSHKDAD 656

Qy       57 -NGKKT----EVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVT----AKEVIDA 107
              | :     |: : :||  |  |           |   : || :  :|    |||  |
Db      657 VIGSQVASGGELSVESKTGNINVK---------AAERQQNIDEQKTALTVNGYAKEAGDK 707

Qy      108 VNKAGWRM-------KTT---------------------TANGQTGQADKFETVTSGTN 138
              :|| |:       |||                     | :|    ||| :   ||
Db      708 QYRAGLRIEHTRDSEKTTRTENSASSLSGGSVKLKAEKDVTFSGSKLVADKGDASVSGNK 767

Qy      139 VTFASGKGTTA-----------------------------------------TVSKDDQ 156
            |:| :    ||                                         |   | :
Db      768 VSFLAADDKTASNTEQTKIGGGFYYTGGIDKLGSGVEAGYENNKTQAQSSKAITSGSDVK 827

Qy      157 GNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAG 216
            ||:|    :|  | |     |:|      :  ||      :   | :  | |  |||
Db      828 GNLT----INARDKLTQQGAQHSVGGAYQENAAGVDHLAAADTASTTTTKTDVGVNI--- 880

Qy      217 NNIEITRNGKNIDIATSMTPQFSSVSLGA-----------GADAPTLSVDGDALNVGSK 264
                  | |:| :    |   :|   |            |  || : :|  |    |:
Db      881 --------GANVDYSAVTRPVERAVGKAAKLDATGVINDIGGIGAPNVGLDIGAQGGSSE 932
```

```
Qy        265 K--DNKPVRITNVAPGV----KEGDV----TNVAQLKGVAQNLN----------NRIDNV 304
              |   :    :::|  |      :|:|    |     || | ||           || |
Db        933 KRSSSSQAVVSSVQAGSIDINAKGEVRDQGTQYQASKG-AVNLTADSHRSEAAANRQDEQ 991

Qy        305 DGNARAGIAQAIATAGLVQAYLPGKSMMAIG----GGTYRGEAGYAIGYSSISDGGNW-- 358
              : |        :|| |: |    | : :     ||| |  :  :   :   |  |
Db        992 SRDTR-------GSAG-VRVYTTTGSDLTVDAKGEGGTQRSNSSASQAVTGSIDAANGIN 1043

Qy        359 -------IIKGTASGNSRGHFGASA 376
                    | :|||    ||    :|
Db       1044 VNVKKDAIYQGTALNGGRGKTAVNA 1068


RESULT 11
Y741_CHLMU
ID   Y741_CHLMU     STANDARD;      PRT;  1007 AA.
AC   Q9PJT6;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Hypothetical protein TC0741 precursor.
GN   TC0741.
OS   Chlamydia muridarum.
OC   Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX   NCBI_TaxID=83560;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MoPn / Nigg;
RX   MEDLINE=20150255; PubMed=10684935;
RA   Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA   White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,
RA   Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA   Gwinn M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
RA   Eisen J., Fraser C.M.;
RT   "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
RT   pneumoniae AR39.";
RL   Nucleic Acids Res. 28:1397-1406(2000).
CC   -!- SIMILARITY: BELONGS TO THE CHLAMYDIAL CPN0572/CT456/TC0741
CC       FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE002342; AAF39550.1; -.
DR   PIR; H81670; H81670.
DR   TIGR; TC0741; -.
KW   Hypothetical protein; Signal; Complete proteome.
FT   SIGNAL        1     51       POTENTIAL.
FT   CHAIN        52   1007       HYPOTHETICAL PROTEIN TC0741.
SQ   SEQUENCE   1007 AA;  104006 MW;  842800C0871B1518 CRC64;

  Query Match             7.1%;  Score 137.5;  DB 1;  Length 1007;
  Best Local Similarity   21.9%;  Pred. No. 1.5;
  Matches  109;  Conservative   52;  Mismatches  165;  Indels  171;  Gaps   26;

Qy          9 VLNAGWNIKGVK-----------PGT-TASDNVDFVRTYDTVEFLSADTKT---TTVNVE 53
              |:|   |: |               || |:| :||   : || :  ::|  |   |  ::
Db        423 VINVNVNVGGTNVNIGDTNVSKGSGTPTSSQSVDM--STDTSDLDTSDIDTNNQTNGDIN 480

Qy         54 SKDNGKKTEVKIGAKTSVIKEKDG----KLVTGKDK---------------------- 85
              : ||    :  :    | ::: ||    :   | |
Db        481 TNDNSNNVDGSLSDVDSRVEDDDGVSDTESTNGNDSGKTTSTEENGDPSGPDILAAVRKH 540

Qy         86 ------GENGSSTDEGEGLVTAKE----VIDAVNKAGWRMKTTTANGQTGQADKFETVTS 135
                   |||| ||   || : | :    ||  | : |   :|    | ||  |   :| :
Db        541 LDTVYPGENGGST---EGPLPANQNLGNVIHDVEQNGSAKETIITPGDTGPTDSSSSVDA 597

Qy        136 GTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKV 195
                :|   |   | : :  ||           | |  : |   |||    ::     ||
Db        598 DADVEDTS--DTDSGIGDDD----------GVSDTESTNG-NNSGKTTSTEENGDPSGPD 644

Qy        196 I------------SGNVSPSKGKMDETVNINAGNNI-EITRNGKNIDIATSMTP----- 236
              |                |   ::| :    | | || | :: :||   :  | :||
Db        645 ILAAVRKHLDTVYPGENGGSTEGPL--PANQNLGNVIHDVEQNGAAQE--TIITPGDTES 700

Qy        237 QFSSVSLGAGADAPTLS-----------VDGDALNVGSKKDNKPV-------------- 270
                :| |: | ||   :|            | :: |      | ||
Db        701 TDTSSSVNANADLEDVSDADSGFGDDDGISDTESTNGNDSGKNTPVGDGGTPSGPDILAA 760

Qy        271 ---RITNVAPGVKEGDVTNVAQLKGVAQNLNNRIDNVDGNA----------RAGIAQAIA 317
                 :  | ||   |           ||| : | :|: |          | |      :
Db        761 VRKHLDTVYPGENGGSTERPLPAN---QNLGDIIHDVEQNGSAKETVVSPYRGGGGNTSS 817

Qy        318 TAGLVQAYLP-------------GK----SMMAIGGGTYRGEAGYAIGYSSISDGGNWI 359
              ||  : ||             ||    |:|  || |             : :|||
Db        818 PIGLA-SLLPATPSTPLMTTPRTNGKAAASSLMIKGGETQA---------KLVKNGGN-- 865

Qy        360 IKGTAS-----GNSRGH 371
              | |  :        |||
Db        866 IPGETTLAELLPRLRGH 882


RESULT 12
FLIC_ECOLI
ID   FLIC_ECOLI     STANDARD;      PRT;   497 AA.
AC   P04949;
DT   13-AUG-1987 (Rel. 05, Created)
DT   13-AUG-1987 (Rel. 05, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Flagellin.
GN   FLIC OR FLAF OR HAG OR B1923.
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12;
RX   MEDLINE=87057066; PubMed=3536885;
RA   Kuwajima G., Asaka J.-I., Fujiwara T., Fujiwara T., Node K., Kondo E.;
RT   "Nucleotide sequence of the hag gene encoding flagellin of
RT   Escherichia coli.";
RL   J. Bacteriol. 168:1479-1483(1986).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12;
RX   MEDLINE=89281489; PubMed=2659972;
RA   Hanafusa T., Sakai A., Tominaga A., Enomoto M.;
RT   "Isolation and characterization of Escherichia coli hag operator
RT   mutants whose hag48 expression has become repressible by a Salmonella
RT   H1 repressor.";
RL   Mol. Gen. Genet. 216:44-50(1989).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12 / MG1655;
RX   MEDLINE=97426617; PubMed=9278503;
RA   Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA   Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA   Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA   Mau B., Shao Y.;
RT   "The complete genome sequence of Escherichia coli K-12.";
RL   Science 277:1453-1474(1997).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12;
RX   MEDLINE=97251358; PubMed=9097040;
RA   Itoh T., Aiba H., Baba T., Fujita K., Hayashi K., Inada T.,
RA   Isono K., Kasai H., Kimura S., Kitakawa M., Kitagawa M.,
RA   Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,
RA   Nakade S., Nakamura Y., Nashimoto H., Nishio Y., Oshima T.,
```

```
RA   Saito N., Sampei G., Seki Y., Sivasundaram S., Tagami H.,
RA   Takeda J., Takemoto K., Wada C., Yamamoto Y., Horiuchi T.;
RT   "A 460-kb DNA sequence of the Escherichia coli K-12 genome
RT   corresponding to the 40.1-50.0 min region on the linkage map.";
RL   DNA Res. 3:379-392(1996).
RN   [5]
RP   SEQUENCE OF 1-19 FROM N.A.
RX   MEDLINE=83238225; PubMed=6305924;
RA   Szekely E., Simon M.;
RT   "DNA sequence adjacent to flagellar genes and evolution of flagellar-
RT   phase variation.";
RL   J. Bacteriol. 155:74-81(1983).
RN   [6]
RP   SEQUENCE OF 1-12.
RC   STRAIN=K12 / EMG2;
RX   MEDLINE=97443975; PubMed=9298646;
RA   Link A.J., Robison K., Church G.M.;
RT   "Comparing the predicted and observed properties of proteins encoded
RT   in the genome of Escherichia coli K-12.";
RL   Electrophoresis 18:1259-1313(1997).
RN   [7]
RP   SEQUENCE OF 1-4.
RC   STRAIN=K12 / W3110;
RX   MEDLINE=98263247; PubMed=9600841;
RA   Wilkins M.R., Gasteiger E., Tonella L., Ou K., Tyler M.,
RA   Sanchez J.-C., Gooley A.A., Walsh B.J., Bairoch A., Appel R.D.,
RA   Williams K.L., Hochstrasser D.F.;
RT   "Protein identification with N and C-terminal sequence tags in
RT   proteome projects.";
RL   J. Mol. Biol. 278:599-608(1998).
CC   -!- FUNCTION: FLAGELLIN IS THE SUBUNIT PROTEIN WHICH POLYMERIZES TO
CC       FORM THE FILAMENTS OF BACTERIAL FLAGELLA.
CC   -!- SIMILARITY: BELONGS TO THE BACTERIAL FLAGELLIN FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; M14358; AAA23950.1; -.
DR   EMBL; X17440; CAA35488.1; -.
DR   EMBL; AE000285; AAC74990.1; -.
DR   EMBL; D90832; BAA15744.1; -.
DR   EMBL; D90833; BAA15751.1; -.
DR   EMBL; J01607; AAA92491.1; -.
DR   PIR; A37249; FLEC.
DR   SWISS-2DPAGE; P04949; COLI.
DR   EcoGene; EG10321; fliC.
DR   InterPro; IPR001029; Flagellin_C.
DR   InterPro; IPR001492; FlagellinN.
DR   Pfam; PF00700; Flagellin_C; 1.
DR   Pfam; PF00669; Flagellin_N; 1.
DR   PRINTS; PR00207; FLAGELLIN.
KW   Flagella; Complete proteome.
FT   INIT_MET      0      0
FT   CONFLICT    284    284       P -> L (IN REF. 2).
SQ   SEQUENCE   497 AA;  51163 MW;  24B5419C21C7B4E8 CRC64;

  Query Match              7.0%;  Score 137;  DB 1;  Length 497;
  Best Local Similarity   20.6%;  Pred. No. 0.71;
  Matches   80;  Conservative   71;  Mismatches  154;  Indels   84;  Gaps   16;

Qy         22 GTTASDNVDFVR-----TYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKD 76
              || :  ::  ::         | :: :|  |:   ||| :|:   |   :::||  :    |
Db        101 GTNSESDLSSIQDEIKSRLDEIDRVSGQTQFNGVNVLAKNGSMK--IQVGANDNQTITID 158

Qy         77 GKLVTGKDKGENGSSTDEGEGLVTAKEV-------IDAVNKAGWRMKT--TTANGQTGQA 127
              | :  |  | :| |     : : |:  |          : :    |  : |    | | |  |
Db        159 LKQIDAKTLGLDGFSVKNNDTVTTSAPVTAFGATTTNNIKLTGITLSTEAATDTGGTNPA 218

Qy        128 -----------DKFETVTSGTN-------------VTFASGKGTTATVSKDD------- 155
                         | :  :| | |                 || |:|    |||:  :
Db        219 SIEGVYTDNGNDYYAKITGGDNDGKYYAVTVANDGTVTMATGATANATVTDANTTKATTI 278

Qy        156 -QGNITVMYDVNVGD------ALNVNQLQNS-GWNLDSKAVAGSSGKVISGNVSPSKGKM 207
                |    |    |    |     |::: :||:| | : |: |:   ::| : : :|: : | :
Db        279 TSGGTPVQIDNTAGSATANLGAVSLVKLQDSKGNDTDTYALKDTNGNLYAADVNETTGAV 338

Qy        208 -------DETVNINAGNNIEIITRNGKNIDIATSMTPQFSSVSLGAGADAPTLSVDGDAL 259
                     | :   ::   :::  :     ::          : |  |  |       |:  |:||
Db        339 SVKTITYTDSSGAASSPTAVKLGGDDGKTEVVDIDGKTYDSADLNGGNLQTGLTAGGEAL 398

Qy        260 N-VGSKKDNKPVRITNVAPGVKE------GDVTNVAQLKGVAQNLNNRIDNVDGNARAGI 312
                | : |    |:: :  : |     :          | | |  :|       ||||   |:    |:: |
Db        399 TAVANGKTTDPLKALDDAIASVDKFRSSLGAVQN--RLDSAVTNLNNTTTNL-SEAQSRI 455

Qy        313 AQA--------IATAGLVQAYLPGKSMMA 333
                |         :: | ::|    | |::|
Db        456 QDADYATEVSNMSKAQIIQQ--AGNSVLA 482


RESULT 13
ICEN_XANCT
ID   ICEN_XANCT     STANDARD;      PRT;  1567 AA.
AC   P18127;
DT   01-NOV-1990 (Rel. 16, Created)
DT   01-NOV-1990 (Rel. 16, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Ice nucleation protein.
GN   INAX.
OS   Xanthomonas campestris (pv. translucens).
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC   Xanthomonadaceae; Xanthomonas.
OX   NCBI_TaxID=343;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=X56S;
RX   MEDLINE=91080859; PubMed=2259339;
RA   Zhao J., Orser C.S.;
RT   "Conserved repetition in the ice nucleation gene inaX from
RT   Xanthomonas campestris pv. translucens.";
RL   Mol. Gen. Genet. 223:163-166(1990).
CC   -!- FUNCTION: ICE NUCLEATION PROTEINS ENABLE BACTERIA TO NUCLEATE
CC       CRYSTALLIZATION IN SUPERCOOLED WATER.
CC   -!- SUBCELLULAR LOCATION: Outer membrane (By similarity).
CC   -!- DOMAIN: CONTAINS 153 IMPERFECT REPEATS OF THE CONSENSUS
CC       OCTAPEPTIDE A-G-Y-G-S-T-L-T; FURTHER ON A 16-RESIDUE AND A
CC       REGIONAL 48-RESIDUE PERIODICITY IS SUPERIMPOSED.
CC   -!- MISCELLANEOUS: A STRUCTURAL MODEL IS SUGGESTED IN WHICH THE ICE
CC       NUCLEATION PROTEIN DISPLAYS A SYMMETRY RELATED TO THAT OF ICE.
CC   -!- SIMILARITY: BELONGS TO THE BACTERIAL ICE NUCLEATION PROTEIN
CC       FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; X52970; CAA37140.1; -.
DR   HSSP; P06620; 1INA.
DR   InterPro; IPR000258; Ice_nucleatn.
DR   Pfam; PF00818; Ice_nucleation; 81.
DR   PRINTS; PR00327; ICENUCLEATN.
DR   PROSITE; PS00314; ICE_NUCLEATION; 57.
KW   Ice nucleation; Repeat; Outer membrane.
SQ   SEQUENCE   1567 AA;  152548 MW;  C8B451D959ECAD63 CRC64;

  Query Match              7.0%;  Score 137;  DB 1;  Length 1567;
  Best Local Similarity   21.9%;  Pred. No. 2.6;
  Matches   85;  Conservative   46;  Mismatches  176;  Indels   82;  Gaps   15;
```

```
Qy          10 LNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKT 69
Db         762 VTAGYGSTG-----TAGADSTLIAGYGSTQTSGSDSSLTAGYGSTQTARKGSDITAGYGS 816

Qy          70 SVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADK 129
Db         817 TGTAGADSTLIAG-----YGSTQTSGS---------DSSLTAGYGSTQTAREGSDVTAGY 862

Qy         130 FETVTSGTNVTFASGKGTTATVSKDDQGNITVMY----------DVNVGDALNVNQLQNS 179
Db         863 GSTGTAGADSTLISGYGSTQTAGSD--SSLTAGYGSTQTARKGSDVTAG---------- 909

Qy         180 GWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFS 239
Db         910 ---YGSTGTAGADSTLIAGYGSTQTSGSDSSLTAGYGST-QTARKG------SDMTAGYG 959

Qy         240 SVSLGAGADAPTLSVDGDALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNN 299
Db         960 STGT-AGADSTLIAGYGSTQTSGSD--------SSLTAGY------------GSTQTARE 998

Qy         300 RIDNVDGNARAGIAQA----IATAGLVQAYLPGKSMMAIGGGTYRGEAGYAI--GYSSIS 353
Db         999 GSDVTAGYGSTGTAGADSTLIAGYGSTQTAGSDSSLTAGYGSTQTARQGSDVTAGYGSTG 1058

Qy         354 DGG--NWIIKGTASGNSRGHFGASASVGY 380
Db        1059 TAGADSTLIAGYGSTQTAGS-DSSLTAGY 1086
```

RESULT 14
PMP8_CHLPN
ID   PMP8_CHLPN   STANDARD;   PRT;   930 AA.
AC   Q9Z393; Q9RB66;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Probable outer membrane protein pmp8 precursor (Polymorphic membrane
DE   protein 8) (Outer membrane protein 11).
GN   PMP8 OR OMP11 OR CPN0446 OR CP0307.
OS   Chlamydia pneumoniae (Chlamydophila pneumoniae).
OC   Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX   NCBI_TaxID=83558;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=VR1310;
RX   MEDLINE=20007584; PubMed=10539856;
RA   Christiansen G., Boesen T., Hjerno K., Daugaard L., Mygind P.,
RA   Madsen A.S., Knudsen K., Falk E., Birkelund S.;
RT   "Molecular biology of Chlamydia pneumoniae surface proteins and their
RT   role in immunopathogenicity.";
RL   Am. Heart J. 138:S491-S495(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CWL029;
RX   MEDLINE=99206606; PubMed=10192388;
RA   Kalman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,
RA   Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
RT   "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";
RL   Nat. Genet. 21:385-389(1999).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AR39;
RX   MEDLINE=20150255; PubMed=10684935;
RA   Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA   White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,
RA   Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA   Gwinn M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
RA   Eisen J., Fraser C.M.;
RT   "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
RT   pneumoniae AR39.";
RL   Nucleic Acids Res. 28:1397-1406(2000).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=J138;
RX   MEDLINE=20330349; PubMed=10871362;
RA   Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
RA   Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
RT   "Comparison of whole genome sequences of Chlamydia pneumoniae J138
RT   from Japan and CWL029 from USA.";
RL   Nucleic Acids Res. 28:2311-2314(2000).
CC   -!- SUBCELLULAR LOCATION: CELL WALL SURFACE (ELEMENTARY BODIES)
CC   (POTENTIAL).
CC   -!- SIMILARITY: BELONGS TO THE PMP OUTER MEMBRANE PROTEIN FAMILY.
CC   ---
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---
DR   EMBL; AJ133034; CAB37068.1; -.
DR   EMBL; AE001627; AAD18590.1; -.
DR   EMBL; AE002193; AAF38164.1; -.
DR   EMBL; AP002546; BAA98654.1; -.
DR   PIR; A81591; A81591.
DR   PIR; D72078; D72078.
DR   PHCI-2DPAGE; Q9Z393; -.
DR   TIGR; CP0307; -.
DR   InterPro; IPR006315; Autotransport.
DR   InterPro; IPR005546; Autotransporter.
DR   InterPro; IPR003368; Chlamydia_PMP.
DR   Pfam; PF03797; Autotransporter; 1.
DR   Pfam; PF02415; DUF145; 2.
DR   TIGRFAMs; TIGR01414; autotrans_barl; 1.
DR   TIGRFAMs; TIGR01376; POMP_repeat; 6.
KW   Outer membrane; Signal; Multigene family; Complete proteome.
FT   SIGNAL   1   26   POTENTIAL.
FT   CHAIN   27   930   PROBABLE OUTER MEMBRANE PROTEIN PMP8.
FT   CONFLICT   177   177   T -> A (IN REF. 3 AND 4).
SQ   SEQUENCE   930 AA; 97669 MW; 46A9B5E3BB913C4C CRC64;

 Query Match   7.0%; Score 136.5; DB 1; Length 930;
 Best Local Similarity   22.6%; Pred. No. 1.6;
 Matches   88; Conservative   48; Mismatches 156; Indels   97; Gaps   18;

```
Qy          37 TVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKDKGE-NGSSTDEG 95
Db          34 TDSFDGAGGSTFTPKSTADANG--TNYVLSGNVYINDAGKGTALTGCCFTETTGDLTFTG 91

Qy          96 EGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKFETVTSGTNVTFASGKGTTATVSKD- 154
Db          92 KGYSFSFNTVDAGSNAGAAASTT--------ADKALTFTGFSNLSFIAAPGTTVASGKST 143

Qy         155 ----------DQGNITVMYDVNV-------GDALNVNQLQNSGWNLDSKAVAGSSGKVIS 197
Db         144 LSSAGALNLTDNG--TILFSQNVSNEANNNGGAITTKTLSISG-NTSSITFTSNSAKKLG 200

Qy         198 GNVSPSKGKMDETVNINAGNNIEITRNGK------NIDIATSMTPQFSSV---------S 242
Db         201 GAIYSSAA---ASISGNTGQLVFMNNKGETGGGALGFEASSSIT-QNSSLFFSGNTATDA 256

Qy         243 LGAG--------ADAPTLSVDGDALNVGSKKDNKPVRITNVAPGVKEGDV-TNVAQLKGV 293
Db         257 AGKGGAIYCEKTGETPTLTISG----------NKSLTFAENSSVTQGGAICAHGLDLSAA 306

Qy         294 AQNL--NNRIDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSS 351
Db         307 GPTLFSNNRC----GNTAAGKGGAIAIAD------------------SGSLSLSANQGD 343

Qy         352 ISDGGNWIIKGTASGNSRG--HFGASASV 378
Db         344 ITFLGNTLTSTSAPTSTRNAIYLGSSAKI 372
```

```
RESULT 15
HLYA_PROMI
ID   HLYA_PROMI     STANDARD;      PRT;  1577 AA.
AC   P16466;
DT   01-AUG-1990 (Rel. 15, Created)
DT   01-AUG-1990 (Rel. 15, Last sequence update)
DT   01-NOV-1990 (Rel. 16, Last annotation update)
DE   Hemolysin precursor.
GN   HPMA.
OS   Proteus mirabilis.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Proteus.
OX   NCBI_TaxID=584;
RN   [1]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 30-43.
RC   STRAIN=Isolate 477-12;
RX   MEDLINE=90170827; PubMed=2407716;
RA   Uphoff T.S., Welch R.A.;
RT   "Nucleotide sequencing of the Proteus mirabilis calcium-independent
RT   hemolysin genes (hpmA and hpmB) reveals sequence similarity with the
RT   Serratia marcescens hemolysin genes (shlA and shlB).";
RL   J. Bacteriol. 172:1206-1216(1990).
CC   -!- FUNCTION: BACTERIAL HEMOLYSINS ARE EXOTOXINS THAT ATTACK BLOOD
CC       CELL MEMBRANES AND CAUSE CELL RUPTURE BY MECHANISMS NOT CLEARLY
CC       DEFINED.
CC   -!- FUNCTION: CELL-BOUND HEMOLYSIN, WHICH RELEASES HEME-IRON FROM
CC       ERYTHROCYTES BY INTERACTION WITH THE ERYTHROCYTE MEMBRANE. HPMA
CC       REQUIRES HPMB FUNCTION.
CC   -!- SUBCELLULAR LOCATION: Outer membrane.
CC   -!- MISCELLANEOUS: THE CONSERVED AMPHIPATHIC DOMAINS IN SHLA AND HPMA
CC       MAY BE RESPONSIBLE FOR PORE FORMATION.
CC   -!- SIMILARITY: TO S.MARCESCENS HEMOLYSIN (SHLA).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M30186; AAA25657.1; -.
DR   PIR; A35140; A35140.
KW   Hemolysis; Toxin; Outer membrane; Signal.
FT   SIGNAL        1     29
FT   CHAIN        30   1577       HEMOLYSIN.
SQ   SEQUENCE   1577 AA;  165869 MW;  175975E0C924B2D9 CRC64;

  Query Match             7.0%;  Score 136;  DB 1;  Length 1577;
  Best Local Similarity   23.0%;  Pred. No. 3;
  Matches   97;  Conservative   58;  Mismatches  141;  Indels  126;  Gaps   25;

Qy          6 VKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLS--ADTKTTTVNVE---------- 53
              ||  :  | |    |||   :|    |   | :  |:  :: :|  | ||
Db        880 VKKAIEDGVNT--TKPGNN-TDLTKKVTARDAIANLANLSNLETPNVGVEVGIKGGGSQQ 936

Qy         54 SKDNGKKTEVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDA-VNKAG 112
             .|: : :     | |    | : : ||    |:| :  || ||  |       :| ::|
Db        937 SQTDSQAVSTSINAGKIDI-DSNNKL---HDQGTHYQSTQEGISLTANTHTSEATLDKHQ 992

Qy        113 WRMKTTTANGQTGQADKFETVTSGTNVTFA-SGKGTTA------TVSKDDQGNITVMYDV 165
                  |   || | : |     :|:::| |  |:| |       | :|  |        :
Db        993 TTFHETKGGGQIGVSTK-----TGSDITVAIKGEGQTTDNALMETKAKGSQFTSNGDISI 1047

Qy        166 NVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITR-- 223
              |||:          |   |::                |||   || |||| :: : :
Db       1048 NVGE-----NAHYEGAQFDAQ----------------KGK---TV-INAGGDLTLAQAT 1081

Qy        224 ----------NGK-NIDIATSMTPQFSSVSLGAGADAPTL--SVDGDALNVGSKKDNKPV 270
                       ||  |: : |  ||:  |   | | :| |   | :     ||:   :: :
Db       1082 DTHSESQSNVNGSANLKVGT--TPE--SKDYGGGFNAGTTHHSKEQTTAKVGTITGSQGI 1137

Qy        271 RITNVAPGVKEGDVTNVAQLKGVAQNLNNRID------------NVDGNARAGIAQ--- 314
              :         :|  |:::  : :| |  |::|             |: |  :||  :
Db       1138 ELNAGHNLTLQG--THLSSEQDIALNATNKVDLQSASSEHTEKGNNLSGGVQAGFGKKMT 1195

Qy        315 --AIATAGLVQAYLPGKSMMAIG----------GGTYRGEAGYAIGYSSISDGGNWIIKG 362
                 | :  ||      | :  |||          |||             |:: ||  | |
Db       1196 DDASSVNGL------GSAQFAIGKQDEKSVSREGGT------------INNSGNLTING 1236

Qy        363 TA 364
               :
Db       1237 NS 1238


Search completed: October  6, 2003, 09:24:20
Job time : 7.89906 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: October 6, 2003, 09:13:10 ; Search time 28.0565 Seconds
(without alignments)
3513.485 Million cell updates/sec

Title: US-09-771-382-38
Perfect score: 1947
Sequence: 1 NRAASVKDVLNAGWNIKGVK.....TASGNSRGHFGASASVGYQW 382

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_23:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1941	99.7	591	2	Q93QY3	Q93qy3 neisseria m
2	1941	99.7	591	2	Q9JPS7	Q9jps7 neisseria m
3	1941	99.7	591	16	Q9JR18	Q9jr18 neisseria m
4	1937	99.5	592	2	Q9AQF0	Q9aqf0 neisseria m
5	1920	98.6	526	2	Q9JPS4	Q9jps4 neisseria m
6	1920	98.6	530	2	Q9JPS1	Q9jps1 neisseria m
7	1903	97.7	600	2	Q9JPS6	Q9jps6 neisseria m
8	1895.5	97.4	592	2	Q9JPS9	Q9jps9 neisseria m
9	1889.5	97.0	590	2	Q9JPS3	Q9jps3 neisseria m
10	1886.5	96.9	594	2	Q93QY4	Q93qy4 neisseria m
11	1883.5	96.7	594	2	Q9JPI3	Q9jpi3 neisseria m
12	1883.5	96.7	594	2	Q9JPS2	Q9jps2 neisseria m
13	1881.5	96.6	599	2	Q9JPR8	Q9jpr8 neisseria m
14	1874.5	96.3	594	2	Q9JPH7	Q9jph7 neisseria m
15	1874.5	96.3	598	2	Q9JPR9	Q9jpr9 neisseria m
16	1861.5	95.6	592	2	Q93QY2	Q93qy2 neisseria m
17	1861.5	95.6	598	2	Q9JPS0	Q9jps0 neisseria m
18	1861.5	95.6	598	2	Q93QY5	Q93qy5 neisseria m
19	1861.5	95.6	598	2	Q9JPT0	Q9jpt0 neisseria m
20	1860.5	95.6	598	2	Q9JPR7	Q9jpr7 neisseria m
21	1860.5	95.6	599	2	Q9JPS8	Q9jps8 neisseria m
22	1854.5	95.2	592	16	Q9JQW4	Q9jqw4 neisseria m
23	1852.5	95.1	595	2	Q9JPH0	Q9jph0 neisseria m
24	1846.5	94.8	600	2	Q9JPS5	Q9jps5 neisseria m
25	1829.5	94.0	589	2	Q9JPI0	Q9jpi0 neisseria m
26	1817.5	93.3	589	2	Q93QY1	Q93qy1 neisseria m
27	710	36.5	1098	2	Q48152	Q48152 haemophilus
28	708.5	36.4	1096	2	Q8GM79	Q8gm79 haemophilus
29	679.5	34.9	1204	2	Q8GM76	Q8gm76 haemophilus
30	662.5	34.0	1210	2	Q8GM74	Q8gm74 haemophilus
31	660	33.9	2353	2	P71401	P71401 haemophilus
32	653.5	33.6	1210	2	Q8GM75	Q8gm75 haemophilus
33	418.5	21.5	1004	2	Q8GM77	Q8gm77 haemophilus
34	417.5	21.4	1002	2	Q8GM78	Q8gm78 haemophilus
35	374.5	19.2	1299	16	Q9F3X6	Q9f3x6 pasteurella
36	352.5	18.1	2314	2	Q8KQM8	Q8kqm8 moraxella c
37	336.5	17.3	2059	16	Q9PD50	Q9pd50 xylella fas
38	335	17.2	1190	16	Q9PC04	Q9pc04 xylella fas
39	325.5	16.7	1588	16	Q8XDG4	Q8xdg4 escherichia
40	320	16.4	1461	16	Q8ZL64	Q8zl64 salmonella
41	318.5	16.4	1778	16	Q8FCB2	Q8fcb2 escherichia
42	318	16.3	1107	16	Q9F2D8	Q9f2d8 salmonella
43	312	16.0	641	16	Q8CKM1	Q8ckm1 yersinia pe
44	312	16.0	658	16	Q8ZHJ0	Q8zhj0 yersinia pe
45	300.5	15.4	2712	16	Q9F3X5	Q9f3x5 pasteurella

ALIGNMENTS

RESULT 1
Q93QY3
ID Q93QY3 PRELIMINARY; PRT; 591 AA.
AC Q93QY3;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE NhhA outer membrane protein.
GN NHHA.
OS Neisseria meningitidis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=487;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EG329;
RA Peak I.R., Srikhanta Y., Dieckelman M., Moxon R., Jennings M.P.;
RT "Identification and characterization of a gene encoding a novel outer
RT membrane protein of Neisseria meningitidis.";
RL Submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF157606; AAK68867.1; -.
DR InterPro; IPR005594; YadA.
DR Pfam; PF03895; YadA; 1.
SQ SEQUENCE 591 AA; 62048 MW; C0DC600798859C65 CRC64;

Query Match 99.7%; Score 1941; DB 2; Length 591;
Best Local Similarity 100.0%; Pred. No. 4.8e-84;
Matches 381; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          2 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKT 61
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        211 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKT 270

Qy         62 EVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        271 EVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 330

Qy        122 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGW 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        331 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGW 390

Qy        182 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        391 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 450

Qy        242 SLGAGADAPTLSVDGDALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRI 301
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        451 SLGAGADAPTLSVDGDALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRI 510

Qy        302 DNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIK 361
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        511 DNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIK 570

Qy        362 GTASGNSRGHFGASASVGYQW 382
              |||||||||||||||||||||
Db        571 GTASGNSRGHFGASASVGYQW 591


RESULT 2
Q9JPS7
ID   Q9JPS7      PRELIMINARY;      PRT;   591 AA.
AC   Q9JPS7;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Outer membrane protein GNA992.
GN   GNA992.
OS   Neisseria meningitidis.
OC   Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC   Neisseriaceae; Neisseria.
OX   NCBI_TaxID=487;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BZ147;
RX   MEDLINE=20175756; PubMed=10710308;
RA   Pizza M., Scarlato V., Masignani V., Giuliani M.M., Arico' B.,
RA   Comanducci M., Jennings G.T., Baldi L., Bartolini E., Capecchi B.,
RA   Galeotti C.L., Luzzi E., Manetti R., Marchetti E., Mora M., Nuti S.,
RA   Ratti G., Santini L., Savino S., Scarselli M., Storni E., Zuo P.,
RA   Broeker M., Hundt E., Knapp B., Blair E., Mason T., Tettelin H.,
RA   Hood D.W., Jeffries A.C., Saunders N.J., Granoff D.M., Venter C.,
RA   Moxon E.R., Grandi G., Rappuoli R.;
RT   "Identification of Vaccine Candidates Against Serogroup B
RT   Meningococcus by Whole-Genome Sequencing.";
RL   Science 287:1816-1820(2000).
DR   EMBL; AF226366; AAF42515.1; -.
DR   InterPro; IPR005594; YadA.
DR   Pfam; PF03895; YadA; 1.
SQ   SEQUENCE   591 AA;  62113 MW;  533453CAE5A91E1F CRC64;

  Query Match            99.7%;  Score 1941;  DB 2;  Length 591;
  Best Local Similarity  100.0%;  Pred. No. 4.8e-84;
  Matches  381;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKT 61
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        211 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKT 270

Qy         62 EVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        271 EVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 330

Qy        122 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGW 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        331 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGW 390

Qy        182 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        391 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 450

Qy        242 SLGAGADAPTLSVDGDALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRI 301
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        451 SLGAGADAPTLSVDGDALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRI 510

Qy        302 DNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIK 361
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        511 DNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIK 570

Qy        362 GTASGNSRGHFGASASVGYQW 382
              |||||||||||||||||||||
Db        571 GTASGNSRGHFGASASVGYQW 591


RESULT 3
Q9JR18
ID   Q9JR18      PRELIMINARY;      PRT;   591 AA.
AC   Q9JR18;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Outer membrane protein GNA992 (Adhesin) (NhhA outer membrane
DE   protein).
GN   GNA992 OR NMB0992 OR NHHA.
OS   Neisseria meningitidis, and
OS   Neisseria meningitidis (serogroup B).
OC   Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC   Neisseriaceae; Neisseria.
OX   NCBI_TaxID=487, 491;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MC58 / Serogroup B, BZ169, BZ83, and H44/76;
RX   MEDLINE=20175756; PubMed=10710308;
RA   Pizza M., Scarlato V., Masignani V., Giuliani M.M., Arico' B.,
RA   Comanducci M., Jennings G.T., Baldi L., Bartolini E., Capecchi B.,
RA   Galeotti C.L., Luzzi E., Manetti R., Marchetti E., Mora M., Nuti S.,
RA   Ratti G., Santini L., Savino S., Scarselli M., Storni E., Zuo P.,
RA   Broeker M., Hundt E., Knapp B., Blair E., Mason T., Tettelin H.,
RA   Hood D.W., Jeffries A.C., Saunders N.J., Granoff D.M., Venter C.,
RA   Moxon E.R., Grandi G., Rappuoli R.;
RT   "Identification of Vaccine Candidates Against Serogroup B
RT   Meningococcus by Whole-Genome Sequencing.";
RL   Science 287:1816-1820(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MC58 / Serogroup B;
RX   MEDLINE=20175755; PubMed=10710307;
RA   Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
RA   Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
RA   Nelson W.C., Gwinn M.L., DeBoy R., Peterson J.D., Hickey E.K.,
RA   Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
RA   Mason T., Ciecko A., Parksey D.S., Blair E., Cittone H., Clark E.B.,
RA   Cotton M.D., Utterback T.R., Khouri H., Qin H., Vamathevan J.,
RA   Gill J., Scarlato V., Masignani V., Pizza M., Grandi G., Sun L.,
RA   Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;
RT   "Complete genome sequence of Neisseria meningitidis serogroup B strain
RT   MC58.";
RL   Science 287:1809-1815(2000).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   SPECIES=N.meningitidis; STRAIN=PMC21;
RA   Peak I.R., Srikhanta Y., Dieckelman M., Moxon R., Jennings M.P.;
RT   "Identification and characterization of a gene encoding a novel outer
RT   membrane protein of Neisseria meningitidis.";
RL   Submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF226375; AAF42524.1; -.
DR   EMBL; AE002450; AAF41395.1; -.
DR   EMBL; AF226367; AAF42516.1; -.
DR   EMBL; AF226370; AAF42519.1; -.
DR   EMBL; AF226374; AAF42523.1; -.
DR   EMBL; AF157611; AAK68872.1; -.
DR   TIGR; NMB0992; -.
DR   InterPro; IPR005594; YadA.
DR   Pfam; PF03895; YadA; 1.
KW   Complete proteome.
SQ   SEQUENCE   591 AA;  62112 MW;  7C22F3CAE7F73EC6 CRC64;
```

```
  Query Match           99.7%;  Score 1941;  DB 16;  Length 591;
  Best Local Similarity  100.0%;  Pred. No. 4.8e-84;
  Matches  381;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKT 61
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        211 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKT 270

Qy         62 EVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        271 EVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 330

Qy        122 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGW 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        331 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGW 390

Qy        182 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        391 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 450

Qy        242 SLGAGADAPTLSVDGDALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRI 301
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        451 SLGAGADAPTLSVDGDALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRI 510

Qy        302 DNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIK 361
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        511 DNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIK 570

Qy        362 GTASGNSRGHFGASASVGYQW 382
              |||||||||||||||||||||
Db        571 GTASGNSRGHFGASASVGYQW 591


RESULT 4
Q9AQF0
ID   Q9AQF0      PRELIMINARY;      PRT;   592 AA.
AC   Q9AQF0;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Outer membrane protein.
GN   NHHA.
OS   Neisseria meningitidis.
OC   Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC   Neisseriaceae; Neisseria.
OX   NCBI_TaxID=487;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MC58;
RA   Peak I.R., Srikhanta Y., Dieckelman M., Moxon E.R., Jennings M.P.;
RT   "Identification and characterisation of a gene encoding a novel outer
RT   membrane protein of Neisseria meningitidis.";
RL   Submitted (FEB-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF125375; AAK09243.1; -.
DR   InterPro; IPR005594; YadA.
DR   Pfam; PF03895; YadA; 1.
SQ   SEQUENCE   592 AA;  62290 MW;  168986A97381EFC5 CRC64;

  Query Match           99.5%;  Score 1937;  DB 2;  Length 592;
  Best Local Similarity  99.7%;  Pred. No. 7.5e-84;
  Matches  380;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKT 61
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        212 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKT 271

Qy         62 EVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 121
              ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        272 EVKIGVKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 331

Qy        122 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGW 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        332 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGW 391

Qy        182 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        392 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 451

Qy        242 SLGAGADAPTLSVDGDALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRI 301
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        452 SLGAGADAPTLSVDGDALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRI 511

Qy        302 DNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIK 361
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        512 DNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIK 571

Qy        362 GTASGNSRGHFGASASVGYQW 382
              |||||||||||||||||||||
Db        572 GTASGNSRGHFGASASVGYQW 592


RESULT 5
Q9JPS4
ID   Q9JPS4      PRELIMINARY;      PRT;   526 AA.
AC   Q9JPS4;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Outer membrane protein GNA992.
GN   GNA992.
OS   Neisseria meningitidis.
OC   Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC   Neisseriaceae; Neisseria.
OX   NCBI_TaxID=487;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NG6/88;
RX   MEDLINE=20175756; PubMed=10710308;
RA   Pizza M., Scarlato V., Masignani V., Giuliani M.M., Arico' B.,
RA   Comanducci M., Jennings G.T., Baldi L., Bartolini E., Capecchi B.,
RA   Galeotti C.L., Luzzi E., Manetti R., Marchetti E., Mora M., Nuti S.,
RA   Ratti G., Santini L., Savino S., Scarselli M., Storni E., Zuo P.,
RA   Broeker M., Hundt E., Knapp B., Blair E., Mason T., Tettelin H.,
RA   Hood D.W., Jeffries A.C., Saunders N.J., Granoff D.M., Venter C.,
RA   Moxon E.R., Grandi G., Rappuoli R.;
RT   "Identification of Vaccine Candidates Against Serogroup B
RT   Meningococcus by Whole-Genome Sequencing.";
RL   Science 287:1816-1820(2000).
DR   EMBL; AF226377; AAF42526.1; -.
DR   InterPro; IPR005594; YadA.
DR   Pfam; PF03895; YadA; 1.
SQ   SEQUENCE   526 AA;  54732 MW;  F543C4234AF807CA CRC64;

  Query Match           98.6%;  Score 1920;  DB 2;  Length 526;
  Best Local Similarity  99.0%;  Pred. No. 4.1e-83;
  Matches  377;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKT 61
              |||||||||||||||||||||||||||||| |||||||||||||||||||||||||:|||
Db        146 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVHTYDTVEFLSADTKTTTVNVESKDNGKRT 205

Qy         62 EVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 121
              |||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db        206 EVKIGAKTSVIKEKDGKLVTGKGKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 265

Qy        122 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGW 181
              |||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||||
Db        266 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVKYDVNVGDALNVNQLQNSGW 325

Qy        182 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        326 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 385

Qy        242 SLGAGADAPTLSVDGDALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRI 301
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        386 SLGAGADAPTLSVDGDALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRI 445

Qy        302 DNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIK 361
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        446 DNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIK 505

Qy        362 GTASGNSRGHFGASASVGYQW 382
              |||||||||||||||||||||
Db        506 GTASGNSRGHFGASASVGYQW 526


RESULT 6
Q9JPS1
ID   Q9JPS1      PRELIMINARY;      PRT;   530 AA.
AC   Q9JPS1;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Outer membrane protein GNA992.
GN   GNA992.
OS   Neisseria meningitidis.
OC   Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC   Neisseriaceae; Neisseria.
OX   NCBI_TaxID=487;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NGF26;
RX   MEDLINE=20175756; PubMed=10710308;
RA   Pizza M., Scarlato V., Masignani V., Giuliani M.M., Arico' B.,
RA   Comanducci M., Jennings G.T., Baldi L., Bartolini E., Capecchi B.,
RA   Galeotti C.L., Luzzi E., Manetti R., Marchetti E., Mora M., Nuti S.,
RA   Ratti G., Santini L., Savino S., Scarselli M., Storni E., Zuo P.,
RA   Broeker M., Hundt E., Knapp B., Blair E., Mason T., Tettelin H.,
RA   Hood D.W., Jeffries A.C., Saunders N.J., Granoff D.M., Venter C.,
RA   Moxon E.R., Grandi G., Rappuoli R.;
RT   "Identification of Vaccine Candidates Against Serogroup B
RT   Meningococcus by Whole-Genome Sequencing.";
RL   Science 287:1816-1820(2000).
DR   EMBL; AF226380; AAF42529.1; -.
DR   InterPro; IPR005594; YadA.
DR   Pfam; PF03895; YadA; 1.
SQ   SEQUENCE   530 AA;  55190 MW;  1F836CA57598515B CRC64;

  Query Match             98.6%;  Score 1920;  DB 2;  Length 530;
  Best Local Similarity   99.0%;  Pred. No. 4.1e-83;
  Matches  377;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKT 61
              |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||:|
Db        150 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVHTYDTVEFLSADTKTTTVNVESKDNGKRT 209

Qy         62 EVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 121
              |||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db        210 EVKIGAKTSVIKEKDGKLVTGKGKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 269

Qy        122 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGW 181
              |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        270 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVKYDVNVGDALNVNQLQNSGW 329

Qy        182 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 241
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        330 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 389

Qy        242 SLGAGADAPTLSVDGDALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRI 301
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        390 SLGAGADAPTLSVDGDALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRI 449

Qy        302 DNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIK 361
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        450 DNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIK 509

Qy        362 GTASGNSRGHFGASASVGYQW 382
              |||||||||||||||||||||
Db        510 GTASGNSRGHFGASASVGYQW 530


RESULT 7
Q9JPS6
ID   Q9JPS6      PRELIMINARY;      PRT;   600 AA.
AC   Q9JPS6;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Outer membrane protein GNA992.
GN   GNA992.
OS   Neisseria meningitidis.
OC   Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC   Neisseriaceae; Neisseria.
OX   NCBI_TaxID=487;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=E26;
RX   MEDLINE=20175756; PubMed=10710308;
RA   Pizza M., Scarlato V., Masignani V., Giuliani M.M., Arico' B.,
RA   Comanducci M., Jennings G.T., Baldi L., Bartolini E., Capecchi B.,
RA   Galeotti C.L., Luzzi E., Manetti R., Marchetti E., Mora M., Nuti S.,
RA   Ratti G., Santini L., Savino S., Scarselli M., Storni E., Zuo P.,
RA   Broeker M., Hundt E., Knapp B., Blair E., Mason T., Tettelin H.,
RA   Hood D.W., Jeffries A.C., Saunders N.J., Granoff D.M., Venter C.,
RA   Moxon E.R., Grandi G., Rappuoli R.;
RT   "Identification of Vaccine Candidates Against Serogroup B
RT   Meningococcus by Whole-Genome Sequencing.";
RL   Science 287:1816-1820(2000).
DR   EMBL; AF226371; AAF42520.1; -.
DR   InterPro; IPR005594; YadA.
DR   Pfam; PF03895; YadA; 1.
SQ   SEQUENCE   600 AA;  62762 MW;  36256963E0598CD1 CRC64;

  Query Match             97.7%;  Score 1903;  DB 2;  Length 600;
  Best Local Similarity   98.4%;  Pred. No. 3e-82;
  Matches  375;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          2 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKT 61
              |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db        220 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVHTYDTVEFLSADTKTTTVNVESKDNGKKT 279

Qy         62 EVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 121
              |||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db        280 EVKIGAKTSVIKEKDGKLVTGKGKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 339

Qy        122 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGW 181
              || |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        340 GQKGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVKYDVNVGDALNVNQLQNSGW 399

Qy        182 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 241
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        400 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 459

Qy        242 SLGAGADAPTLSVDGDALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRI 301
              |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        460 SLGAGADAPTLSVDGDALNVGSKDANKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNRI 519

Qy        302 DNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIK 361
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        520 DNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWIIK 579

Qy        362 GTASGNSRGHFGASASVGYQW 382
              |||||||||||||||||||||
Db        580 GTASGNSRGHFGASASVGYQW 600


RESULT 8
Q9JPS9
ID   Q9JPS9      PRELIMINARY;      PRT;   592 AA.
AC   Q9JPS9;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
```

```
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Outer membrane protein GNA992.
GN   GNA992.
OS   Neisseria meningitidis.
OC   Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC   Neisseriaceae; Neisseria.
OX   NCBI_TaxID=487;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=860800;
RX   MEDLINE=20175756; PubMed=10710308;
RA   Pizza M., Scarlato V., Masignani V., Giuliani M.M., Arico' B.,
RA   Comanducci M., Jennings G.T., Baldi L., Bartolini E., Capecchi B.,
RA   Galeotti C.L., Luzzi E., Manetti R., Marchetti E., Mora M., Nuti S.,
RA   Ratti G., Santini L., Savino S., Scarselli M., Storni E., Zuo P.,
RA   Broeker M., Hundt E., Knapp B., Blair E., Mason T., Tettelin H.,
RA   Hood D.W., Jeffries A.C., Saunders N.J., Granoff D.M., Venter C.,
RA   Moxon E.R., Grandi G., Rappuoli R.;
RT   "Identification of Vaccine Candidates Against Serogroup B
RT   Meningococcus by Whole-Genome Sequencing.";
RL   Science 287:1816-1820(2000).
DR   EMBL; AF226361; AAF42510.1; -.
DR   InterPro; IPR005594; YadA.
DR   Pfam; PF03895; YadA; 1.
SQ   SEQUENCE   592 AA;  61917 MW;  4A3471514FD3C879 CRC64;

  Query Match            97.4%;  Score 1895.5;  DB 2;  Length 592;
  Best Local Similarity  98.4%;  Pred. No. 6.7e-82;
  Matches  376;  Conservative    0;  Mismatches    5;  Indels    1;  Gaps    1;

Qy          2 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKT 61
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        211 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKT 270

Qy         62 EVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 121
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        271 EVKIGAKTSVIKEKDGKLVTGKGKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 330

Qy        122 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGW 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        331 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGW 390

Qy        182 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        391 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 450

Qy        242 SLGAGADAPTLSVDG-DALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNR 300
              |||||||||||||   |||||||  ||||||||||||||||||||||||||||||||||||
Db        451 SLGAGADAPTLSVDDKGALNVGSKDANKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNR 510

Qy        301 IDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWII 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        511 IDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWII 570

Qy        361 KGTASGNSRGHFGASASVGYQW 382
              ||||||||||||||||||||||
Db        571 KGTASGNSRGHFGASASVGYQW 592


RESULT 9
Q9JPS3
ID   Q9JPS3      PRELIMINARY;      PRT;   590 AA.
AC   Q9JPS3;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Outer membrane protein GNA992.
GN   GNA992.
OS   Neisseria meningitidis.
OC   Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC   Neisseriaceae; Neisseria.
OX   NCBI_TaxID=487;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NGE28;
RX   MEDLINE=20175756; PubMed=10710308;
RA   Pizza M., Scarlato V., Masignani V., Giuliani M.M., Arico' B.,
RA   Comanducci M., Jennings G.T., Baldi L., Bartolini E., Capecchi B.,
RA   Galeotti C.L., Luzzi E., Manetti R., Marchetti E., Mora M., Nuti S.,
RA   Ratti G., Santini L., Savino S., Scarselli M., Storni E., Zuo P.,
RA   Broeker M., Hundt E., Knapp B., Blair E., Mason T., Tettelin H.,
RA   Hood D.W., Jeffries A.C., Saunders N.J., Granoff D.M., Venter C.,
RA   Moxon E.R., Grandi G., Rappuoli R.;
RT   "Identification of Vaccine Candidates Against Serogroup B
RT   Meningococcus by Whole-Genome Sequencing.";
RL   Science 287:1816-1820(2000).
DR   EMBL; AF226378; AAF42527.1; -.
DR   InterPro; IPR005594; YadA.
DR   Pfam; PF03895; YadA; 1.
SQ   SEQUENCE   590 AA;  61661 MW;  8AA476AC300D80C8 CRC64;

  Query Match            97.0%;  Score 1889.5;  DB 2;  Length 590;
  Best Local Similarity  97.9%;  Pred. No. 1.3e-81;
  Matches  374;  Conservative    2;  Mismatches    5;  Indels    1;  Gaps    1;

Qy          2 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKT 61
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:|
Db        209 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKRT 268

Qy         62 EVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 121
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        269 EVKIGAKTSVIKEKDGKLVTGKGKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 328

Qy        122 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGW 181
              ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        329 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVKYDVNVGDALNVNQLQNSGW 388

Qy        182 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        389 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 448

Qy        242 SLGAGADAPTLSVDGD-ALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNR 300
              |||||||||||||| : |||||||  ||||||||||||||||||||||||||||||||||||
Db        449 SLGAGADAPTLSVDDEGALNVGSKDANKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNR 508

Qy        301 IDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWII 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        509 IDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWII 568

Qy        361 KGTASGNSRGHFGASASVGYQW 382
              ||||||||||||||||||||||
Db        569 KGTASGNSRGHFGASASVGYQW 590


RESULT 10
Q93QY4
ID   Q93QY4      PRELIMINARY;      PRT;   594 AA.
AC   Q93QY4;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   NhhA outer membrane protein.
GN   NHHA.
OS   Neisseria meningitidis.
OC   Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC   Neisseriaceae; Neisseria.
OX   NCBI_TaxID=487;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=EG327;
RA   Peak I.R., Srikhanta Y., Dieckelman M., Moxon R., Jennings M.P.;
RT   "Identification and characterization of a gene encoding a novel outer
RT   membrane protein of Neisseria meningitidis.";
RL   Submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF157605; AAK68866.1; -.
```

```
DR   InterPro; IPR005594; YadA.
DR   Pfam; PF03895; YadA; 1.
SQ   SEQUENCE   594 AA;  62297 MW;  9DDD48B04B3A8EA2 CRC64;

  Query Match             96.9%;  Score 1886.5;  DB 2;  Length 594;
  Best Local Similarity   97.6%;  Pred. No. 1.8e-81;
  Matches  373;  Conservative    3;  Mismatches    5;  Indels    1;  Gaps    1;

Qy          2 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKT 61
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:|
Db        213 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKRT 272

Qy         62 EVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 121
              |||||||||||||||||||||||||| ||||:|||||||||||||||||||||||||||||
Db        273 EVKIGAKTSVIKEKDGKLVTGKDKGENDSSTDKGEGLVTAKEVIDAVNKAGWRMKTTTAN 332

Qy        122 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGW 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        333 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGW 392

Qy        182 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        393 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 452

Qy        242 SLGAGADAPTLSVDGD-ALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNR 300
              |||||||||||||| : |||||||  ||||||||||||||||||||||||||||||||||
Db        453 SLGAGADAPTLSVDDEGALNVGSKDANKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNH 512

Qy        301 IDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWII 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        513 IDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWII 572

Qy        361 KGTASGNSRGHFGASASVGYQW 382
              ||||||||||||||||||||||
Db        573 KGTASGNSRGHFGASASVGYQW 594


RESULT 11
Q9JPI3
ID   Q9JPI3      PRELIMINARY;      PRT;   594 AA.
AC   Q9JPI3;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Outer membrane protein GNA992.
GN   GNA992.
OS   Neisseria meningitidis.
OC   Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC   Neisseriaceae; Neisseria.
OX   NCBI_TaxID=487;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NG3/88, and BZ232;
RX   MEDLINE=20175756; PubMed=10710308;
RA   Pizza M., Scarlato V., Masignani V., Giuliani M.M., Arico' B.,
RA   Comanducci M., Jennings G.T., Baldi L., Bartolini E., Capecchi B.,
RA   Galeotti C.L., Luzzi E., Manetti R., Marchetti E., Mora M., Nuti S.,
RA   Ratti G., Santini L., Savino S., Scarselli M., Storni E., Zuo P.,
RA   Broeker M., Hundt E., Knapp B., Blair E., Mason T., Tettelin H.,
RA   Hood D.W., Jeffries A.C., Saunders N.J., Granoff D.M., Venter C.,
RA   Moxon E.R., Grandi G., Rappuoli R.;
RT   "Identification of Vaccine Candidates Against Serogroup B
RT   Meningococcus by Whole-Genome Sequencing.";
RL   Science 287:1816-1820(2000).
DR   EMBL; AF226376; AAF42525.1; -.
DR   EMBL; AF226369; AAF42518.1; -.
DR   InterPro; IPR005594; YadA.
DR   Pfam; PF03895; YadA; 1.
SQ   SEQUENCE   594 AA;  62086 MW;  1B25E03B90D04B46 CRC64;

  Query Match             96.7%;  Score 1883.5;  DB 2;  Length 594;
  Best Local Similarity   97.6%;  Pred. No. 2.5e-81;
  Matches  373;  Conservative    1;  Mismatches    7;  Indels    1;  Gaps    1;

Qy          2 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKT 61
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        213 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKT 272

Qy         62 EVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        273 EVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 332

Qy        122 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGW 181
              ||||||||||||||||| |||||| ||||||||||||||| ||||||||||||||||||||
Db        333 GQTGQADKFETVTSGTKVTFASGNGTTATVSKDDQGNITVKYDVNVGDALNVNQLQNSGW 392

Qy        182 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        393 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 452

Qy        242 SLGAGADAPTLSVDGD-ALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNR 300
              |||||||||||||| : |||||||  ||||||||||||||||||||||||||||||||||
Db        453 SLGAGADAPTLSVDDEGALNVGSKDANKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNH 512

Qy        301 IDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWII 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        513 IDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWII 572

Qy        361 KGTASGNSRGHFGASASVGYQW 382
              ||||||||||||||||||||||
Db        573 KGTASGNSRGHFGASASVGYQW.594


RESULT 12
Q9JPS2
ID   Q9JPS2      PRELIMINARY;      PRT;   594 AA.
AC   Q9JPS2;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Outer membrane protein GNA992.
GN   GNA992.
OS   Neisseria meningitidis.
OC   Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC   Neisseriaceae; Neisseria.
OX   NCBI_TaxID=487;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NGE31;
RX   MEDLINE=20175756; PubMed=10710308;
RA   Pizza M., Scarlato V., Masignani V., Giuliani M.M., Arico' B.,
RA   Comanducci M., Jennings G.T., Baldi L., Bartolini E., Capecchi B.,
RA   Galeotti C.L., Luzzi E., Manetti R., Marchetti E., Mora M., Nuti S.,
RA   Ratti G., Santini L., Savino S., Scarselli M., Storni E., Zuo P.,
RA   Broeker M., Hundt E., Knapp B., Blair E., Mason T., Tettelin H.,
RA   Hood D.W., Jeffries A.C., Saunders N.J., Granoff D.M., Venter C.,
RA   Moxon E.R., Grandi G., Rappuoli R.;
RT   "Identification of Vaccine Candidates Against Serogroup B
RT   Meningococcus by Whole-Genome Sequencing.";
RL   Science 287:1816-1820(2000).
DR   EMBL; AF226379; AAF42528.1; -.
DR   InterPro; IPR005594; YadA.
DR   Pfam; PF03895; YadA; 1.
SQ   SEQUENCE   594 AA;  62114 MW;  1E2A63A78F53D256 CRC64;

  Query Match             96.7%;  Score 1883.5;  DB 2;  Length 594;
  Best Local Similarity   97.6%;  Pred. No. 2.5e-81;
  Matches  373;  Conservative    1;  Mismatches    7;  Indels    1;  Gaps    1;

Qy          2 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKT 61
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        213 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKT 272

Qy         62 EVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        273 EVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 332
```

```
Qy        122 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGW 181
              |||||||||||||||| |||||| |||||||||||||||| ||||||||||||||||||||
Db        333 GQTGQADKFETVTSGTKVTFASGNGTTATVSKDDQGNITVKYDVNVGDALNVNQLQNSGW 392

Qy        182 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 241
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        393 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 452

Qy        242 SLGAGADAPTLSVDGD-ALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNR 300
              |||||||||||||| : ||||||| ||||||||||||||||||||||||||||||||||||
Db        453 SLGAGADAPTLSVDDEGALNVGSKDANKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNH 512

Qy        301 IDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWII 360
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        513 IDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWII 572

Qy        361 KGTASGNSRGHFGASASVGYQW 382
              ||||||||||||||||||||||
Db        573 KGTASGNSRGHFGASASVGYQW 594
```

RESULT 13
Q9JPR8
ID Q9JPR8 PRELIMINARY; PRT; 599 AA.
AC Q9JPR8;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Outer membrane protein GNA992 (NhhA outer membrane protein).
GN GNA992 OR NHHA.
OS Neisseria meningitidis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=487;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NGH38;
RX MEDLINE=20175756; PubMed=10710308;
RA Pizza M., Scarlato V., Masignani V., Giuliani M.M., Arico' B.,
RA Comanducci M., Jennings G.T., Baldi L., Bartolini E., Capecchi B.,
RA Galeotti C.L., Luzzi E., Manetti R., Marchetti E., Mora M., Nuti S.,
RA Ratti G., Santini L., Savino S., Scarselli M., Storni E., Zuo P.,
RA Broeker M., Hundt E., Knapp B., Blair E., Mason T., Tettelin H.,
RA Hood D.W., Jeffries A.C., Saunders N.J., Granoff D.M., Venter C.,
RA Moxon E.R., Grandi G., Rappuoli R.;
RT "Identification of Vaccine Candidates Against Serogroup B
RT Meningococcus by Whole-Genome Sequencing.";
RL Science 287:1816-1820(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=H38;
RA Peak I.R., Srikhanta Y., Dieckelman M., Moxon R., Jennings M.P.;
RT "Identification and characterization of a gene encoding a novel outer
RT membrane protein of Neisseria meningitidis.";
RL Submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF226383; AAF42532.1; -.
DR EMBL; AF157608; AAK68869.1; -.
DR InterPro; IPR005594; YadA.
DR Pfam; PF03895; YadA; 1.
SQ SEQUENCE 599 AA; 62844 MW; BBA16EBF53C1970C CRC64;

Query Match 96.6%; Score 1881.5; DB 2; Length 599;
Best Local Similarity 97.6%; Pred. No. 3.1e-81;
Matches 373; Conservative 1; Mismatches 7; Indels 1; Gaps 1;

```
Qy          2 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKT 61
              |||||||||||||||||||||||||||||| |||||||||||||||||||||||||:|
Db        218 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVHTYDTVEFLSADTKTTTVNVESKDNGKRT 277

Qy         62 EVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 121
              |||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||
Db        278 EVKIGAKTSVIKEKDGKLVTGKGKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 337
```

```
Qy        122 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGW 181
              |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        338 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVKYDVNVGDALNVNQLQNSGW 397

Qy        182 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 241
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        398 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 457

Qy        242 SLGAGADAPTLSVDG-DALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNR 300
              |||||||||||||    |||||||  ||||||||||||||||||||||||||||||||||||
Db        458 SLGAGADAPTLSVDDKGALNVGSKDANKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNR 517

Qy        301 IDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWII 360
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        518 IDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWII 577

Qy        361 KGTASGNSRGHFGASASVGYQW 382
              ||||||||||||||||||||||
Db        578 KGTASGNSRGHFGASASVGYQW 599
```

RESULT 14
Q9JPH7
ID Q9JPH7 PRELIMINARY; PRT; 594 AA.
AC Q9JPH7;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Outer membrane protein GNA992 (NhhA outer membrane protein).
GN GNA992 OR NHHA.
OS Neisseria meningitidis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=487;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BZ198, and 297-0;
RX MEDLINE=20175756; PubMed=10710308;
RA Pizza M., Scarlato V., Masignani V., Giuliani M.M., Arico' B.,
RA Comanducci M., Jennings G.T., Baldi L., Bartolini E., Capecchi B.,
RA Galeotti C.L., Luzzi E., Manetti R., Marchetti E., Mora M., Nuti S.,
RA Ratti G., Santini L., Savino S., Scarselli M., Storni E., Zuo P.,
RA Broeker M., Hundt E., Knapp B., Blair E., Mason T., Tettelin H.,
RA Hood D.W., Jeffries A.C., Saunders N.J., Granoff D.M., Venter C.,
RA Moxon E.R., Grandi G., Rappuoli R.;
RT "Identification of Vaccine Candidates Against Serogroup B
RT Meningococcus by Whole-Genome Sequencing.";
RL Science 287:1816-1820(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BZ198;
RA Peak I.R., Srikhanta Y., Dieckelman M., Moxon R., Jennings M.P.;
RT "Identification and characterization of a gene encoding a novel outer
RT membrane protein of Neisseria meningitidis.";
RL Submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF226368; AAF42517.1; -.
DR EMBL; AF226358; AAF42507.1; -.
DR EMBL; AF157604; AAK68865.1; -.
DR InterPro; IPR005594; YadA.
DR Pfam; PF03895; YadA; 1.
SQ SEQUENCE 594 AA; 62361 MW; 436BDDED68263C5C CRC64;

Query Match 96.3%; Score 1874.5; DB 2; Length 594;
Best Local Similarity 97.4%; Pred. No. 6.6e-81;
Matches 372; Conservative 1; Mismatches 8; Indels 1; Gaps 1;

```
Qy          2 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKT 61
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        213 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKT 272

Qy         62 EVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 121
              |||||||||||||||||||||| | |||||||||||||||||||||||||||||||||||||
```

```
Db        273 EVKIGAKTSVIKEKDGKLVTGKGKDENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 332

Qy        122 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGW 181
              ||||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db        333 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVKYDVNVGDALNVNQLQNSGW 392

Qy        182 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        393 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMAPQFSSV 452

Qy        242 SLGAGADAPTLSVDGD-ALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNR 300
              |||||||||||||| : ||||||| |||||||||||||||||||||||||||||||||||
Db        453 SLGAGADAPTLSVDDEGALNVGSKDTNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNR 512

Qy        301 IDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWII 360
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        513 IDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGDTYRGEAGYAIGYSSISDGGNWII 572

Qy        361 KGTASGNSRGHFGASASVGYQW 382
              ||||||||||||||||||||||
Db        573 KGTASGNSRGHFGASASVGYQW 594
```

```
RESULT 15
Q9JPR9
ID   Q9JPR9      PRELIMINARY;      PRT;   598 AA.
AC   Q9JPR9;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Outer membrane protein GNA992.
GN   GNA992.
OS   Neisseria meningitidis.
OC   Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC   Neisseriaceae; Neisseria.
OX   NCBI_TaxID=487;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NGH36;
RX   MEDLINE=20175756; PubMed=10710308;
RA   Pizza M., Scarlato V., Masignani V., Giuliani M.M., Arico' B.,
RA   Comanducci M., Jennings G.T., Baldi L., Bartolini E., Capecchi B.,
RA   Galeotti C.L., Luzzi E., Manetti R., Marchetti E., Mora M., Nuti S.,
RA   Ratti G., Santini L., Savino S., Scarselli M., Storni E., Zuo P.,
RA   Broeker M., Hundt E., Knapp B., Blair E., Mason T., Tettelin H.,
RA   Hood D.W., Jeffries A.C., Saunders N.J., Granoff D.M., Venter C.,
RA   Moxon E.R., Grandi G., Rappuoli R.;
RT   "Identification of Vaccine Candidates Against Serogroup B
RT   Meningococcus by Whole-Genome Sequencing.";
RL   Science 287:1816-1820(2000).
DR   EMBL; AF226382; AAF42531.1; -.
DR   InterPro; IPR005594; YadA.
DR   Pfam; PF03895; YadA; 1.
SQ   SEQUENCE   598 AA;  62718 MW;  9095F8E31AD7C76D CRC64;

  Query Match             96.3%;  Score 1874.5;  DB 2;  Length 598;
  Best Local Similarity   97.4%;  Pred. No. 6.6e-81;
  Matches  372;  Conservative    1;  Mismatches    8;  Indels    1;  Gaps    1;

Qy          2 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKT 61
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        217 RAASVKDVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKT 276

Qy         62 EVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 121
              ||||||||||||||||||||| | ||||||||||||||||||||||||||||||||||||
Db        277 EVKIGAKTSVIKEKDGKLVTGKGKDENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 336

Qy        122 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVMYDVNVGDALNVNQLQNSGW 181
              ||||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db        337 GQTGQADKFETVTSGTNVTFASGKGTTATVSKDDQGNITVKYDVNVGDALNVNQLQNSGW 396

Qy        182 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMTPQFSSV 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        397 NLDSKAVAGSSGKVISGNVSPSKGKMDETVNINAGNNIEITRNGKNIDIATSMAPQFSSV 456

Qy        242 SLGAGADAPTLSVDGD-ALNVGSKKDNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNR 300
              |||||||||||||| : ||||||| |||||||||||||||||||||||||||||||||||
Db        457 SLGAGADAPTLSVDDEGALNVGSKDTNKPVRITNVAPGVKEGDVTNVAQLKGVAQNLNNR 516

Qy        301 IDNVDGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWII 360
              |||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        517 IDNVDGNALAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIGYSSISDGGNWII 576

Qy        361 KGTASGNSRGHFGASASVGYQW 382
              ||||||||||||||||||||||
Db        577 KGTASGNSRGHFGASASVGYQW 598
```

Search completed: October 6, 2003, 09:30:54
Job time : 29.0565 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: October 6, 2003, 09:05:45 ; Search time 17.261 Seconds
(without alignments)
1848.329 Million cell updates/sec

Title: US-09-771-382-39
Perfect score: 1018
Sequence: 1 SANTLKAGDNLKIKQFTYSL.....AGWRMKTTTANGQTGQADKF 201

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03:*
1: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1018	100.0	502	22	AAU06186	N. meningitidis PM
2	957.5	94.1	591	20	AAY27202	Amino acid sequenc
3	957.5	94.1	591	20	AAY23741	A surface protein
4	957.5	94.1	591	20	AAY23746	A surface protein
5	957.5	94.1	591	22	AAU06171	N. meningitidis PM
6	957.5	94.1	591	22	AAU06175	N. meningitidis EG
7	954.5	93.8	591	21	AAY57045	BASB029 amino acid
8	953.5	93.7	592	20	AAY23737	A surface protein
9	925.5	90.9	598	20	AAY23738	A surface protein
10	925.5	90.9	598	22	AAU06178	N. meningitidis BZ
11	925.5	90.9	599	20	AAY23743	A surface protein
12	925.5	90.9	599	22	AAU06176	N. meningitidis H3
13	923.5	90.7	594	20	AAY23739	A surface protein
14	923.5	90.7	594	22	AAU06179	N. meningitidis BZ
15	922.5	90.6	594	20	AAY23740	A surface protein
16	922.5	90.6	594	21	AAY57044	BASB029 amino acid
17	922.5	90.6	594	22	AAU06174	N. meningitidis EG
18	921.5	90.5	598	20	AAY23742	A surface protein
19	921.5	90.5	598	22	AAU06177	N. meningitidis H1
20	918.5	90.2	589	20	AAY23745	A surface protein
21	918.5	90.2	589	22	AAU06173	N. meningitidis P2
22	907.5	89.1	592	20	AAY23744	A surface protein
23	907.5	89.1	592	22	AAU06172	N. meningitidis H4
24	903.5	88.8	592	22	AAU06180	N. meningitidis Z2
25	898.5	88.3	592	20	AAY27203	Amino acid sequenc
26	864.5	84.9	512	22	AAU06182	N. meningitidis PM
27	831.5	81.7	604	22	AAU06181	N. meningitidis su
28	820.5	80.6	513	22	AAU06183	N. meningitidis H4
29	673.5	66.2	245	20	AAY27201	Amino acid sequenc
30	622.5	61.1	2353	17	AAR99393	Haemophilus adhesi
31	622.5	61.1	2411	21	AAB23860	Haemophilus influe
32	618	60.7	1098	17	AAR99392	Haemophilus adhesi
33	608.5	59.8	433	22	AAU06185	N. meningitidis PM
34	555	54.5	1094	21	AAB23858	Haemophilus influe
35	523.5	51.4	407	22	AAU06184	N. meningitidis PM
36	417	41.0	679	17	AAR99394	Haemophilus adhesi
37	417	41.0	679	21	AAB23855	Haemophilus influe
38	374	36.7	298	24	AAE30477	Haemophilus influe
39	327.5	32.2	1004	21	AAB23857	Haemophilus influe
40	320	31.4	1002	21	AAB23854	Haemophilus influe
41	311.5	30.6	1104	21	AAB23856	Haemophilus influe
42	311.5	30.6	1104	21	AAB23859	Haemophilus influe
43	279	27.4	56	21	AAB37827	Neisserial conserv
44	238	23.4	53	21	AAB37828	Neisserial conserv
45	201	19.7	2139	24	ABP71294	M. catarrhalis sur

ALIGNMENTS

RESULT 1
AAU06186
ID AAU06186 standard; Protein; 502 AA.
XX
AC AAU06186;
XX
DT 24-OCT-2001 (first entry)
XX
DE N. meningitidis PMC21 NhhA deletion mutant #4.
XX
KW Surface antigen NhhA; meningococcal disease; meningitis vaccine;
KW mutant; mutein.
XX
OS Neisseria meningitidis strain PMC21.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Peptide 1..49
FT /label= Signal_peptide
FT Protein 50..502
FT /label= Mature_NhhA_deletion_mutant_#4
FT /note= "Predicted mature protein, specifically
FT claimed in claim 12"
XX
PN WO200155182-A1.
XX
PD 02-AUG-2001.
XX
PF 25-JAN-2001; 2001WO-AU00069.
XX
PR 25-JAN-2000; 2000US-0177917.
XX

```
PA   (UYQU ) UNIV QUEENSLAND.
XX
PI   Peak IRA, Jennings MP;
XX
DR   WPI; 2001-488774/53.
DR   N-PSDB; AAS09176.
XX
PT   New NhhA surface antigen polypeptides and polynucleotides from
PT   Neisseria meningitidis, useful in producing vaccines for treating or
PT   preventing broad spectrum of Neisseria meningitidis -
XX
PS   Claim 12; Fig 9; 91pp; English.
XX
CC   The present invention relates to the isolation of novel Neisseria
CC   meningitidis mutant polypeptides of the surface antigen NhhA
CC   (AAU06182-AAU06186). The modified or mutant NhhA polypeptides are
CC   characterised by deletions of non-conserved amino acids, particularly
CC   the deletion of variable regions. The deletion mutants are useful in
CC   diagnostics, therapeutic and prophylactic vaccines against a broader
CC   spectrum of N. meningitidis, and in designing and/or screening of
CC   medicaments. The mutant proteins when used as a vaccine can effectively
CC   immunise against a broader spectrum of N. meningitidis strains than
CC   would be expected from a corresponding wild-type surface antigen.
CC   The present sequence represents N. meningitidis strain PMC21 surface
CC   antigen NhhA deletion mutant #4.
XX
SQ   Sequence   502 AA;

  Query Match             100.0%;  Score 1018;  DB 22;  Length 502;
  Best Local Similarity   100.0%;  Pred. No. 5e-78;
  Matches  201;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 SANTLKAGDNLKIKQFTYSLKKDLTDLTSVGTEKLSFSANGNKVNITSDTKGLNFAKETA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         50 SANTLKAGDNLKIKQFTYSLKKDLTDLTSVGTEKLSFSANGNKVNITSDTKGLNFAKETA 109

Qy         61 GTNGDTTVHLNGIGSTLTDRAASVKDVLNAGWNIKGVKNVDFVRTYDTVEFLSADTKTTT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        110 GTNGDTTVHLNGIGSTLTDRAASVKDVLNAGWNIKGVKNVDFVRTYDTVEFLSADTKTTT 169

Qy        121 VNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVN 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        170 VNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVN 229

Qy        181 KAGWRMKTTTANGQTGQADKF 201
              |||||||||||||||||||||
Db        230 KAGWRMKTTTANGQTGQADKF 250


RESULT 2
AAY27202
ID   AAY27202 standard; Protein; 591 AA.
XX
AC   AAY27202;
XX
DT   24-SEP-1999  (first entry)
XX
DE   Amino acid sequence of N. meningitidis protein ORF40-1.
XX
KW   Neisseria meningitidis protein; pharmaceutical; vaccine; diagnosis;
KW   bacterial infection; treatment.
XX
OS   Neisseria meningitidis.
XX
PN   WO9936544-A2.
XX
PD   22-JUL-1999.
XX
PF   14-JAN-1999;   99WO-IB00103.
XX
PR   09-OCT-1998;   98GB-0022143.
PR   14-JAN-1998;   98GB-0000760.
PR   01-SEP-1998;   98GB-0019015.
XX
PA   (CHIR-) CHIRON SPA.
XX
PI   Grandi G, Masignani V, Pizza M, Rappuoli R, Scarlato V;
XX
DR   WPI; 1999-444400/37.
DR   N-PSDB; AAX99124.
XX
PT   New protein and its nucleotide sequence, useful in vaccines or
PT   diagnostic compositions for treating and/or preventing Neisseria
PT   meningitidis infections
XX
PS   Claim 1; Page 62; 123pp; English.
XX
CC   The invention provides proteins (AAY27201-245) from Neisseria
CC   meningitidis (strains A and B) and nucleic acid sequences (AAX99123-167)
CC   encoding the proteins. Compositions comprising the protein, nucleic acid
CC   or antibody specific to the protein are useful as pharmaceuticals, e.g. a
CC   vaccine composition or a diagnostic composition. The composition is also
CC   useful for treating or preventing an infection due to Neisserial
CC   bacteria, especially Neisseria meningitidis.
XX
SQ   Sequence   591 AA;

  Query Match             94.1%;  Score 957.5;  DB 20;  Length 591;
  Best Local Similarity   85.7%;  Pred. No. 8.3e-73;
  Matches  198;  Conservative    0;  Mismatches    0;  Indels   33;  Gaps    3;

Qy          4 TLKAGDNLKIKQ----FTYSLKKDLTDLTSVGTEKLSFSANGNKVNITSDTKGLNFAKET 59
              ||||||||||||    ||||||||||||||||||||||||||||||||||||||||||||
Db        109 TLKAGDNLKIKQNGTNFTYSLKKDLTDLTSVGTEKLSFSANGNKVNITSDTKGLNFAKET 168

Qy         60 AGTNGDTTVHLNGIGSTLTD----------------------RAASVKDVLNAGWNIKGV 97
              ||||||||||||||||||||                      ||||||||||||||||||
Db        169 AGTNGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGV 228

Qy         98 K-------NVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKL 150
              |       ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        229 KPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKL 288

Qy        151 VTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKF 201
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        289 VTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKF 339


RESULT 3
AAY23741
ID   AAY23741 standard; Protein; 591 AA.
XX
AC   AAY23741;
XX
DT   08-SEP-1999  (first entry)
XX
DE   A surface protein of Neisseria meningitidis.
XX
KW   Surface protein; surface glycoprotein; infection; vaccine;
KW   immunoreactive peptide.
XX
OS   Neisseria meningitidis.
XX
PN   WO9931132-A1.
XX
PD   24-JUN-1999.
XX
PF   14-DEC-1998;   98WO-AU01031.
XX
PR   12-DEC-1997;   97GB-0026398.
XX
PA   (ISIS-) ISIS INNOVATION LTD.
PA   (UYQU ) UNIV QUEENSLAND.
XX
PI   Jennings MP, Moxon ER, Peak IRA;
XX
```

```
DR   WPI; 1999-418754/35.
DR   N-PSDB; AAX85793.
XX
PT   Neisseria meningitidis surface proteins useful for treating N.
PT   meningitidis infections
XX
PS   Claim 1; Page 104-106; 132pp; English.
XX
CC   The present sequence represents a surface protein of Neiserria
CC   meningitidis which is approximately 62 kDa. The N. meningitidis
CC   surface glycoproteins, nucleic acids, the primers and optionally
CC   a thermostable polymerase, or antibodies are useful in a kit for
CC   the detection or diagnosis of N. meningitidis infection in humans.
CC   The N. meningitidis surface glycoproteins can also be used to
CC   prevent or treat N. meningitidis infection in humans, especially
CC   in the form of vaccines. The proteins and antibodies can also
CC   be used to identify immunoreactive peptides.
XX
SQ   Sequence   591 AA;

  Query Match            94.1%;  Score 957.5;  DB 20;  Length 591;
  Best Local Similarity  85.7%;  Pred. No. 8.3e-73;
  Matches  198;  Conservative    0;  Mismatches    0;  Indels   33;  Gaps    3;

Qy          4 TLKAGDNLKIKQ----FTYSLKKDLTDLTSVGTEKLSFSANGNKVNITSDTKGLNFAKET 59
              ||||||||||||    ||||||||||||||||||||||||||||||||||||||||||||
Db        109 TLKAGDNLKIKQNGTNFTYSLKKDLTDLTSVGTEKLSFSANGNKVNITSDTKGLNFAKET 168

Qy         60 AGTNGDTTVHLNGIGSTLTD----------------------RAASVKDVLNAGWNIKGV 97
              ||||||||||||||||||||                      ||||||||||||||||||
Db        169 AGTNGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGV 228

Qy         98 K-------NVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKL 150
              |       ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        229 KPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKL 288

Qy        151 VTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKF 201
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        289 VTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKF 339


RESULT 4
AAY23746
ID   AAY23746 standard; Protein; 591 AA.
XX
AC   AAY23746;
XX
DT   08-SEP-1999  (first entry)
XX
DE   A surface protein of Neisseria meningitidis.
XX
KW   Surface protein; surface glycoprotein; infection; vaccine;
KW   immunoreactive peptide.
XX
OS   Neisseria meningitidis.
XX
PN   WO9931132-A1.
XX
PD   24-JUN-1999.
XX
PF   14-DEC-1998;   98WO-AU01031.
XX
PR   12-DEC-1997;   97GB-0026398.
XX
PA   (ISIS-) ISIS INNOVATION LTD.
PA   (UYQU ) UNIV QUEENSLAND.
XX
PI   Jennings MP,  Moxon ER,  Peak IRA;
XX
DR   WPI; 1999-418754/35.
DR   N-PSDB; AAX85798.
XX
PT   Neisseria meningitidis surface proteins useful for treating N.
PT   meningitidis infections
XX
PS   Claim 1; Page 127-128; 132pp; English.
XX
CC   The present sequence represents a surface protein of Neiserria
CC   meningitidis which is approximately 62 kDa. The N. meningitidis
CC   surface glycoproteins, nucleic acids, the primers and optionally
CC   a thermostable polymerase, or antibodies are useful in a kit for
CC   the detection or diagnosis of N. meningitidis infection in humans.
CC   The N. meningitidis surface glycoproteins can also be used to
CC   prevent or treat N. meningitidis infection in humans, especially
CC   in the form of vaccines. The proteins and antibodies can also
CC   be used to identify immunoreactive peptides.
XX
SQ   Sequence   591 AA;

  Query Match            94.1%;  Score 957.5;  DB 20;  Length 591;
  Best Local Similarity  85.7%;  Pred. No. 8.3e-73;
  Matches  198;  Conservative    0;  Mismatches    0;  Indels   33;  Gaps    3;

Qy          4 TLKAGDNLKIKQ----FTYSLKKDLTDLTSVGTEKLSFSANGNKVNITSDTKGLNFAKET 59
              ||||||||||||    ||||||||||||||||||||||||||||||||||||||||||||
Db        109 TLKAGDNLKIKQNGTNFTYSLKKDLTDLTSVGTEKLSFSANGNKVNITSDTKGLNFAKET 168

Qy         60 AGTNGDTTVHLNGIGSTLTD----------------------RAASVKDVLNAGWNIKGV 97
              ||||||||||||||||||||                      ||||||||||||||||||
Db        169 AGTNGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGV 228

Qy         98 K-------NVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKL 150
              |       ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        229 KPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKL 288

Qy        151 VTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKF 201
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        289 VTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKF 339


RESULT 5
AAU06171
ID   AAU06171 standard; Protein; 591 AA.
XX
AC   AAU06171;
XX
DT   24-OCT-2001  (first entry)
XX
DE   N. meningitidis PMC21 surface antigen NhhA polypeptide sequence.
XX
KW   Surface antigen NhhA; meningococcal disease; meningitis vaccine.
XX
OS   Neisseria meningitidis strain PMC21.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..51
FT                   /label= Signal_peptide
FT   Region          1..50
FT                   /label= C1
FT                   /note= "Conserved region 1"
FT   Region          51..108
FT                   /label= V1
FT                   /note= "Variable region 1"
FT   Protein         52..591
FT                   /label= Mature_NhhA
FT                   /note= "Predicted mature protein, specifically
FT                     claimed in claim 12"
FT   Region          109..120
FT                   /label= C2
FT                   /note= "Conserved region 2"
FT   Region          121..124
FT                   /label= V2
FT                   /note= "Variable region 2"
FT   Region          125..188
FT                   /label= C3
FT                   /note= "Conserved region 3"
```

```
FT   Region          189..210
FT                   /label= V3
FT                   /note= "Variable region 3"
FT   Region          211..229
FT                   /label= C4
FT                   /note= "Conserved region 4"
FT   Region          230..236
FT                   /label= V4
FT                   /note= "Variable region 4"
FT   Region          237..591
FT                   /label= C5
FT                   /note= "Conserved region 5"
XX
PN   WO200155182-A1.
XX
PD   02-AUG-2001.
XX
PF   25-JAN-2001; 2001WO-AU00069.
XX
PR   25-JAN-2000; 2000US-0177917.
XX
PA   (UYQU ) UNIV QUEENSLAND.
XX
PI   Peak IRA,  Jennings MP;
XX
DR   WPI; 2001-488774/53.
DR   N-PSDB; AAS09161.
XX
PT   New NhhA surface antigen polypeptides and polynucleotides from
PT   Neisseria meningitidis, useful in producing vaccines for treating or
PT   preventing broad spectrum of Neisseria meningitidis -
XX
PS   Claim 9; Fig 1; 91pp; English.
XX
CC   The present invention relates to the isolation of novel Neisseria
CC   meningitidis mutant polypeptides of the surface antigen NhhA
CC   (AAU06182-AAU06186). The modified or mutant NhhA polypeptides are
CC   characterised by deletions of non-conserved amino acids, particularly
CC   the deletion of variable regions. The deletion mutants are useful in
CC   diagnostics, therapeutic and prophylactic vaccines against a broader
CC   spectrum of N. meningitidis, and in designing and/or screening of
CC   medicaments. The mutant proteins when used as a vaccine can effectively
CC   immunise against a broader spectrum of N. meningitidis strains than
CC   would be expected from a corresponding wild-type surface antigen.
CC   The present sequence representing the wild type surface antigen NhhA
CC   from N. meningitidis strain PMC21 is 1 of 10 NhhA polypeptide sequences
CC   (AAU06171-AAU06180) from 10 different N. meningitidis strains given in
CC   the present invention.
XX
SQ   Sequence   591 AA;

  Query Match             94.1%;  Score 957.5;  DB 22;  Length 591;
  Best Local Similarity   85.7%;  Pred. No. 8.3e-73;
  Matches  198;  Conservative    0;  Mismatches    0;  Indels   33;  Gaps    3;

Qy          4 TLKAGDNLKIKQ----FTYSLKKDLTDLTSVGTEKLSFSANGNKVNITSDTKGLNFAKET 59
              ||||||||||||    ||||||||||||||||||||||||||||||||||||||||||||
Db        109 TLKAGDNLKIKQNGTNFTYSLKKDLTDLTSVGTEKLSFSANGNKVNITSDTKGLNFAKET 168

Qy         60 AGTNGDTTVHLNGIGSTLTD----------------------RAASVKDVLNAGWNIKGV 97
              ||||||||||||||||||||                      ||||||||||||||||||
Db        169 AGTNGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGV 228

Qy         98 K-------NVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKL 150
              |       ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        229 KPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKL 288

Qy        151 VTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKF 201
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        289 VTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKF 339


RESULT 6
```

```
AAU06175
ID   AAU06175 standard; Protein; 591 AA.
XX
AC   AAU06175;
XX
DT   24-OCT-2001  (first entry)
XX
DE   N. meningitidis EG329 surface antigen NhhA polypeptide sequence.
XX
KW   Surface antigen NhhA; meningococcal disease; meningitis vaccine.
XX
OS   Neisseria meningitidis strain EG329.
XX
FH   Key             Location/Qualifiers
FT   Region          1..50
FT                   /label= C1
FT                   /note= "Conserved region 1"
FT   Region          51..108
FT                   /label= V1
FT                   /note= "Variable region 1"
FT   Region          109..120
FT                   /label= C2
FT                   /note= "Conserved region 2"
FT   Region          121..124
FT                   /label= V2
FT                   /note= "Variable region 2"
FT   Region          125..188
FT                   /label= C3
FT                   /note= "Conserved region 3"
FT   Region          189..210
FT                   /label= V3
FT                   /note= "Variable region 3"
FT   Region          211..229
FT                   /label= C4
FT                   /note= "Conserved region 4"
FT   Region          230..236
FT                   /label= V4
FT                   /note= "Variable region 4"
FT   Region          237..591
FT                   /label= C5
FT                   /note= "Conserved region 5"
XX
PN   WO200155182-A1.
XX
PD   02-AUG-2001.
XX
PF   25-JAN-2001; 2001WO-AU00069.
XX
PR   25-JAN-2000; 2000US-0177917.
XX
PA   (UYQU ) UNIV QUEENSLAND.
XX
PI   Peak IRA,  Jennings MP;
XX
DR   WPI; 2001-488774/53.
DR   N-PSDB; AAS09165.
XX
PT   New NhhA surface antigen polypeptides and polynucleotides from
PT   Neisseria meningitidis, useful in producing vaccines for treating or
PT   preventing broad spectrum of Neisseria meningitidis -
XX
PS   Claim 9; Fig 1; 91pp; English.
XX
CC   The present invention relates to the isolation of novel Neisseria
CC   meningitidis mutant polypeptides of the surface antigen NhhA
CC   (AAU06182-AAU06186). The modified or mutant NhhA polypeptides are
CC   characterised by deletions of non-conserved amino acids, particularly
CC   the deletion of variable regions. The deletion mutants are useful in
CC   diagnostics, therapeutic and prophylactic vaccines against a broader
CC   spectrum of N. meningitidis, and in designing and/or screening of
CC   medicaments. The mutant proteins when used as a vaccine can effectively
CC   immunise against a broader spectrum of N. meningitidis strains than
CC   would be expected from a corresponding wild-type surface antigen.
```

```
CC   The present sequence representing the wild type surface antigen NhhA
CC   from N. meningitidis strain EG329 is 1 of 10 NhhA polypeptide sequences
CC   (AAU06171-AAU06180) from 10 different N. meningitidis strains given in
CC   the present invention.
XX
SQ   Sequence   591 AA;

  Query Match             94.1%;  Score 957.5;  DB 22;  Length 591;
  Best Local Similarity   85.7%;  Pred. No. 8.3e-73;
  Matches  198;  Conservative    0;  Mismatches    0;  Indels   33;  Gaps    3;

Qy          4 TLKAGDNLKIKQ----FTYSLKKDLTDLTSVGTEKLSFSANGNKVNITSDTKGLNFAKET 59
              ||||||||||||    ||||||||||||||||||||||||||||||||||||||||||||
Db        109 TLKAGDNLKIKQNGTNFTYSLKKDLTDLTSVGTEKLSFSANGNKVNITSDTKGLNFAKET 168

Qy         60 AGTNGDTTVHLNGIGSTLTD----------------------RAASVKDVLNAGWNIKGV 97
              ||||||||||||||||||||                      ||||||||||||||||||
Db        169 AGTNGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGV 228

Qy         98 K-------NVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKL 150
              |       ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        229 KPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKL 288

Qy        151 VTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKF 201
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        289 VTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKF 339


RESULT 7
AAY57045
ID   AAY57045 standard; Protein; 591 AA.
XX
AC   AAY57045;
XX
DT   21-FEB-2000  (first entry)
XX
DE   BASB029 amino acid sequence from N. meningitidis strain H44/76.
XX
KW   BASB029; Nisseria meningitidis; surface fibril protein; HSF; diagnosis;
KW   infection; treatment; prevent; antibacterial drug.
XX
OS   Neisseria meningitidis.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 90
FT                   /note= "Encoded by AAT"
FT   Misc-difference 92
FT                   /note= "Encoded by GAT"
FT   Misc-difference 98
FT                   /note= "Encoded by AAC"
FT   Misc-difference 108
FT                   /note= "Encoded by AATC"
FT   Misc-difference 123
FT                   /note= "Encoded by ACA"
FT   Misc-difference 269
FT                   /note= "Encoded by AAA"
FT   Misc-difference 389
FT                   /note= "Encoded by CGT"
XX
PN   WO9958683-A2.
XX
PD   18-NOV-1999.
XX
PF   07-MAY-1999;   99WO-EP03255.
XX
PR   13-MAY-1998;   98GB-0010276.
XX
PA   (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PI   Ruelle J;
XX
DR   WPI; 2000-053103/04.
DR   N-PSDB; AAZ39865.
XX
PT   New polypeptide from neisseria meningitidis useful for diagnosis,
PT   treatment or prevention of bacterial infections in mammal  -
XX
PS   Claim 4; Fig 2; 74pp; English.
XX
CC   This is the Nisseria meningitidis BASB029 amino acid sequence from
CC   serogroup B strain H44/76. The BASB029 protein is homologous to the
CC   Haemophilus influenzae surface fibril (HSF) protein. The invention
CC   relates to BASB029 polynucleotide sequences (AAZ39864-Z39865) and
CC   polypeptide sequences (AAY57044-Y57045) and their immunogenic fragments.
CC   BASB029 polypeptides are useful in a method of diagnosing a Neisseria
CC   meningitidis infection in a mammal. Compositions containing BASB029
CC   polynucleotides and polypeptides are useful for generating an immune
CC   response in an animal. A therapeutic composition comprising an antibody
CC   directed against BAsB029 is useful in treating humans with Neisseria
CC   meningitidis disease. The polynucleotide is useful in the diagnosis of
CC   the stage of infection, type of infection, susceptibility to an
CC   infection which results from increased or decreased expression of the
CC   polynucleotide, and for therapeutic or prophylactic purposes,
CC   particularly genetic immunisation. Antibodies against BASB029
CC   polynucleotides and polypeptides are also useful for treating infections
CC   particularly bacterial infections. The protein is useful in the
CC   screening and development of antibacterial drugs. Fused recombinant
CC   protein is useful for the stimulation of the immune system of an organism
CC   receiving the protein.
XX
SQ   Sequence   591 AA;

  Query Match             93.8%;  Score 954.5;  DB 21;  Length 591;
  Best Local Similarity   85.3%;  Pred. No. 1.5e-72;
  Matches  197;  Conservative    1;  Mismatches    0;  Indels   33;  Gaps    3;

Qy          4 TLKAGDNLKIKQ----FTYSLKKDLTDLTSVGTEKLSFSANGNKVNITSDTKGLNFAKET 59
              ||||||||||||    ||||||||||||||||||||||||||||||||||||||||||||
Db        109 TLKAGDNLKIKQNGSNFTYSLKKDLTDLTSVGTEKLSFSANGNKVNITSDTKGLNFAKET 168

Qy         60 AGTNGDTTVHLNGIGSTLTD----------------------RAASVKDVLNAGWNIKGV 97
              ||||||||||||||||||||                      ||||||||||||||||||
Db        169 AGTNGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGV 228

Qy         98 K-------NVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKL 150
              |       ||||||||||||||||||||||||||||||:||||||||||||||||||||||
Db        229 KPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKRTEVKIGAKTSVIKEKDGKL 288

Qy        151 VTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKF 201
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        289 VTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKF 339


RESULT 8
AAY23737
ID   AAY23737 standard; Protein; 592 AA.
XX
AC   AAY23737;
XX
DT   08-SEP-1999  (first entry)
XX
DE   A surface protein of Neisseria meningitidis.
XX
KW   Surface protein; surface glycoprotein; infection; vaccine;
KW   immunoreactive peptide.
XX
OS   Neisseria meningitidis.
XX
PN   WO9931132-A1.
XX
PD   24-JUN-1999.
XX
PF   14-DEC-1998;   98WO-AU01031.
XX
PR   12-DEC-1997;   97GB-0026398.
XX
```

```
PA   (ISIS-) ISIS INNOVATION LTD.
PA   (UYQU ) UNIV QUEENSLAND.
XX
PI   Jennings MP,  Moxon ER,  Peak IRA;
XX
DR   WPI; 1999-418754/35.
DR   N-PSDB; AAX85788.
XX
PT   Neisseria meningitidis surface proteins useful for treating N.
PT   meningitidis infections
XX
PS   Claim 1; Page 86-87; 132pp; English.
XX
CC   The present sequence represents a surface protein of Neiserria
CC   meningitidis which is approximately 62 kDa. The N. meningitidis
CC   surface glycoproteins, nucleic acids, the primers and optionally
CC   a thermostable polymerase, or antibodies are useful in a kit for
CC   the detection or diagnosis of N. meningitidis infection in humans.
CC   The N. meningitidis surface glycoproteins can also be used to
CC   prevent or treat N. meningitidis infection in humans, especially
CC   in the form of vaccines. The proteins and antibodies can also
CC   be used to identify immunoreactive peptides.
XX
SQ   Sequence   592 AA;

  Query Match             93.7%;  Score 953.5;  DB 20;  Length 592;
  Best Local Similarity   85.3%;  Pred. No. 1.8e-72;
  Matches  197;  Conservative    0;  Mismatches    1;  Indels   33;  Gaps    3;

Qy          4 TLKAGDNLKIKQ----FTYSLKKDLTDLTSVGTEKLSFSANGNKVNITSDTKGLNFAKET 59
              ||||||||||||    ||||||||||||||||||||||||||||||||||||||||||||
Db        110 TLKAGDNLKIKQNGTNFTYSLKKDLTDLTSVGTEKLSFSANGNKVNITSDTKGLNFAKET 169

Qy         60 AGTNGDTTVHLNGIGSTLTD----------------------RAASVKDVLNAGWNIKGV 97
              ||||||||||||||||||||                      ||||||||||||||||||
Db        170 AGTNGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGV 229

Qy         98 K--------NVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKL 150
              |        |||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db        230 KPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGVKTSVIKEKDGKL 289

Qy        151 VTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKF 201
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        290 VTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKF 340


RESULT 9
AAY23738
ID   AAY23738 standard; Protein; 598 AA.
XX
AC   AAY23738;
XX
DT   08-SEP-1999  (first entry)
XX
DE   A surface protein of Neisseria meningitidis.
XX
KW   Surface protein; surface glycoprotein; infection; vaccine;
KW   immunoreactive peptide.
XX
OS   Neisseria meningitidis.
XX
PN   WO9931132-A1.
XX
PD   24-JUN-1999.
XX
PF   14-DEC-1998;   98WO-AU01031.
XX
PR   12-DEC-1997;   97GB-0026398.
XX
PA   (ISIS-) ISIS INNOVATION LTD.
PA   (UYQU ) UNIV QUEENSLAND.
XX
PI   Jennings MP,  Moxon ER,  Peak IRA;
XX
DR   WPI; 1999-418754/35.
DR   N-PSDB; AAX85790.
XX
PT   Neisseria meningitidis surface proteins useful for treating N.
PT   meningitidis infections
XX
PS   Claim 1; Page 91-93; 132pp; English.
XX
CC   The present sequence represents a surface protein of Neiserria
CC   meningitidis which is approximately 62 kDa. The N. meningitidis
CC   surface glycoproteins, nucleic acids, the primers and optionally
CC   a thermostable polymerase, or antibodies are useful in a kit for
CC   the detection or diagnosis of N. meningitidis infection in humans.
CC   The N. meningitidis surface glycoproteins can also be used to
CC   prevent or treat N. meningitidis infection in humans, especially
CC   in the form of vaccines. The proteins and antibodies can also
CC   be used to identify immunoreactive peptides.
XX
SQ   Sequence   598 AA;

  Query Match             90.9%;  Score 925.5;  DB 20;  Length 598;
  Best Local Similarity   79.8%;  Pred. No. 4.3e-70;
  Matches  194;  Conservative    1;  Mismatches    5;  Indels   43;  Gaps    3;

Qy          2 ANTLKAGDNLKIKQ----------FTYSLKKDLTDLTSVGTEKLSFSANGNKVNIT 47
              | ||||||||||||          ||||||||||||||| |||||| |||||||||
Db        103 AITLKAGDNLKIKQNTNENTNENTNDSSFTYSLKKDLTDLTSVETEKLSFGANGNKVNIT 162

Qy         48 SDTKGLNFAKETAGTNGDTTVHLNGIGSTLTD----------------------RAASVK 85
              ||||||||||||||||| ||||||||||||                      ||||||
Db        163 SDTKGLNFAKETAGTNGDPTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVK 222

Qy         86 DVLNAGWNIKGVK-------NVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGA 138
              ||||||||||||       ||||||||||||||||||||||||||||||:|||||||
Db        223 DVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKRTEVKIGA 282

Qy        139 KTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQA 198
              ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        283 KTSVIKEKDGKLVTGKGKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQA 342

Qy        199 DKF 201
              |||
Db        343 DKF 345


RESULT 10
AAU06178
ID   AAU06178 standard; Protein; 598 AA.
XX
AC   AAU06178;
XX
DT   24-OCT-2001  (first entry)
XX
DE   N. meningitidis BZ10 surface antigen NhhA polypeptide sequence.
XX
KW   Surface antigen NhhA; meningococcal disease; meningitis vaccine.
XX
OS   Neisseria meningitidis strain BZ10.
XX
FH   Key             Location/Qualifiers
FT   Region          1..50
FT                   /label= C1
FT                   /note= "Conserved region 1"
FT   Region          51..104
FT                   /label= V1
FT                   /note= "Variable region 1"
FT   Region          105..116
FT                   /label= C2
FT                   /note= "Conserved region 2"
FT   Region          117..130
FT                   /label= V2
FT                   /note= "Variable region 2"
```

```
FT   Region          131..194
FT                   /label= C3
FT                   /note= "Conserved region 3"
FT   Region          195..216
FT                   /label= V3
FT                   /note= "Variable region 3"
FT   Region          217..235
FT                   /label= C4
FT                   /note= "Conserved region 4"
FT   Region          236..242
FT                   /label= V4
FT                   /note= "Variable region 4"
FT   Region          243..598
FT                   /label= C5
FT                   /note= "Conserved region 5"
XX
PN   WO200155182-A1.
XX
PD   02-AUG-2001.
XX
PF   25-JAN-2001; 2001WO-AU00069.
XX
PR   25-JAN-2000; 2000US-0177917.
XX
PA   (UYQU ) UNIV QUEENSLAND.
XX
PI   Peak IRA,  Jennings MP;
XX
DR   WPI; 2001-488774/53.
DR   N-PSDB; AAS09168.
XX
PT   New NhhA surface antigen polypeptides and polynucleotides from
PT   Neisseria meningitidis, useful in producing vaccines for treating or
PT   preventing broad spectrum of Neisseria meningitidis -
XX
PS   Claim 9; Fig 1; 91pp; English.
XX
CC   The present invention relates to the isolation of novel Neisseria
CC   meningitidis mutant polypeptides of the surface antigen NhhA
CC   (AAU06182-AAU06186). The modified or mutant NhhA polypeptides are
CC   characterised by deletions of non-conserved amino acids, particularly
CC   the deletion of variable regions. The deletion mutants are useful in
CC   diagnostics, therapeutic and prophylactic vaccines against a broader
CC   spectrum of N. meningitidis, and in designing and/or screening of
CC   medicaments. The mutant proteins when used as a vaccine can effectively
CC   immunise against a broader spectrum of N. meningitidis strains than
CC   would be expected from a corresponding wild-type surface antigen.
CC   The present sequence representing the wild type surface antigen NhhA
CC   from N. meningitidis strain BZ10 is 1 of 10 NhhA polypeptide sequences
CC   (AAU06171-AAU06180) from 10 different N. meningitidis strains given in
CC   the present invention.
XX
SQ   Sequence   598 AA;

  Query Match            90.9%;  Score 925.5;  DB 22;  Length 598;
  Best Local Similarity   79.8%;  Pred. No. 4.3e-70;
  Matches  194;  Conservative    1;  Mismatches    5;  Indels   43;  Gaps    3;

Qy          2 ANTLKAGDNLKIKQ--------------FTYSLKKDLTDLTSVGTEKLSFSANGNKVNIT 47
              | ||||||||||||               ||||||||||||||| |||||| |||||||||
Db        103 AITLKAGDNLKIKQNTNENTNENTNDSSFTYSLKKDLTDLTSVETEKLSFGANGNKVNIT 162

Qy         48 SDTKGLNFAKETAGTNGDTTVHLNGIGSTLTD----------------------RAASVK 85
              |||||||||||||||| ||||||||||||                        ||||||
Db        163 SDTKGLNFAKETAGTNGDPTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVK 222

Qy         86 DVLNAGWNIKGVK-------NVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGA 138
              ||||||||||||       ||||||||||||||||||||||||||||:|||||||
Db        223 DVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKRTEVKIGA 282

Qy        139 KTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQA 198
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        283 KTSVIKEKDGKLVTGKGKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQA 342

Qy        199 DKF 201
              |||
Db        343 DKF 345


RESULT 11
AAY23743
ID   AAY23743 standard; Protein; 599 AA.
XX
AC   AAY23743;
XX
DT   08-SEP-1999  (first entry)
XX
DE   A surface protein of Neisseria meningitidis.
XX
KW   Surface protein; surface glycoprotein; infection; vaccine;
KW   immunoreactive peptide.
XX
OS   Neisseria meningitidis.
XX
PN   WO9931132-A1.
XX
PD   24-JUN-1999.
XX
PF   14-DEC-1998;   98WO-AU01031.
XX
PR   12-DEC-1997;   97GB-0026398.
XX
PA   (ISIS-) ISIS INNOVATION LTD.
PA   (UYQU ) UNIV QUEENSLAND.
XX
PI   Jennings MP,  Moxon ER,  Peak IRA;
XX
DR   WPI; 1999-418754/35.
DR   N-PSDB; AAX85795.
XX
PT   Neisseria meningitidis surface proteins useful for treating N.
PT   meningitidis infections
XX
PS   Claim 1; Page 114-115; 132pp; English.
XX
CC   The present sequence represents a surface protein of Neiserria
CC   meningitidis which is approximately 62 kDa. The N. meningitidis
CC   surface glycoproteins, nucleic acids, the primers and optionally
CC   a thermostable polymerase, or antibodies are useful in a kit for
CC   the detection or diagnosis of N. meningitidis infection in humans.
CC   The N. meningitidis surface glycoproteins can also be used to
CC   prevent or treat N. meningitidis infection in humans, especially
CC   in the form of vaccines. The proteins and antibodies can also
CC   be used to identify immunoreactive peptides.
XX
SQ   Sequence   599 AA;

  Query Match            90.9%;  Score 925.5;  DB 20;  Length 599;
  Best Local Similarity   80.1%;  Pred. No. 4.3e-70;
  Matches  193;  Conservative    1;  Mismatches    4;  Indels   43;  Gaps    3;

Qy          4 TLKAGDNLKIKQ--------------FTYSLKKDLTDLTSVGTEKLSFSANGNKVNITSD 49
              ||||||||||||               |||||||||||||| |||||| ||||||||||
Db        106 TLKAGDNLKIKQNTNKNTNENTNDSSFTYSLKKDLTDLTSVETEKLSFGANGNKVNITSD 165

Qy         50 TKGLNFAKETAGTNGDTTVHLNGIGSTLTD----------------------RAASVKDV 87
              ||||||||||||||||||||||||||||||                      ||||||||
Db        166 TKGLNFAKETAGTNGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDKKKRAASVKDV 225

Qy         88 LNAGWNIKGVK-------NVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKT 140
              ||||||||||       ||||| ||||||||||||||||||||||:||||||||||
Db        226 LNAGWNIKGVKPGTTASDNVDFVHTYDTVEFLSADTKTTTVNVESKDNGKRTEVKIGAKT 285

Qy        141 SVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADK 200
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        286 SVIKEKDGKLVTGKGKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADK 345
```

```
Qy        201 F 201
              |
Db        346 F 346


RESULT 12
AAU06176
ID   AAU06176 standard; Protein; 599 AA.
XX
AC   AAU06176;
XX
DT   24-OCT-2001  (first entry)
XX
DE   N. meningitidis H38 surface antigen NhhA polypeptide sequence.
XX
KW   Surface antigen NhhA; meningococcal disease; meningitis vaccine.
XX
OS   Neisseria meningitidis strain H38.
XX
FH   Key             Location/Qualifiers
FT   Region          1..50
FT                   /label= C1
FT                   /note= "Conserved region 1"
FT   Region          51..105
FT                   /label= V1
FT                   /note= "Variable region 1"
FT   Region          106..117
FT                   /label= C2
FT                   /note= "Conserved region 2"
FT   Region          118..131
FT                   /label= V2
FT                   /note= "Variable region 2"
FT   Region          132..195
FT                   /label= C3
FT                   /note= "Conserved region 3"
FT   Region          196..217
FT                   /label= V3
FT                   /note= "Variable region 3"
FT   Region          218..236
FT                   /label= C4
FT                   /note= "Conserved region 4"
FT   Region          237..243
FT                   /label= V4
FT                   /note= "Variable region 4"
FT   Region          244..599
FT                   /label= C5
FT                   /note= "Conserved region 5"
XX
PN   WO200155182-A1.
XX
PD   02-AUG-2001.
XX
PF   25-JAN-2001; 2001WO-AU00069.
XX
PR   25-JAN-2000; 2000US-0177917.
XX
PA   (UYQU ) UNIV QUEENSLAND.
XX
PI   Peak IRA,  Jennings MP;
XX
DR   WPI; 2001-488774/53.
DR   N-PSDB; AAS09166.
XX
PT   New NhhA surface antigen polypeptides and polynucleotides from
PT   Neisseria meningitidis, useful in producing vaccines for treating or
PT   preventing broad spectrum of Neisseria meningitidis -
XX
PS   Claim 9; Fig 1; 91pp; English.
XX
CC   The present invention relates to the isolation of novel Neisseria
CC   meningitidis mutant polypeptides of the surface antigen NhhA
CC   (AAU06182-AAU06186). The modified or mutant NhhA polypeptides are
CC   characterised by deletions of non-conserved amino acids, particularly
CC   the deletion of variable regions. The deletion mutants are useful in
CC   diagnostics, therapeutic and prophylactic vaccines against a broader
CC   spectrum of N. meningitidis, and in designing and/or screening of
CC   medicaments. The mutant proteins when used as a vaccine can effectively
CC   immunise against a broader spectrum of N. meningitidis strains than
CC   would be expected from a corresponding wild-type surface antigen.
CC   The present sequence representing the wild type surface antigen NhhA
CC   from N. meningitidis strain H38 is 1 of 10 NhhA polypeptide sequences
CC   (AAU06171-AAU06180) from 10 different N. meningitidis strains given in
CC   the present invention.
XX
SQ   Sequence   599 AA;

  Query Match           90.9%;  Score 925.5;  DB 22;  Length 599;
  Best Local Similarity   80.1%;  Pred. No. 4.3e-70;
  Matches  193;  Conservative    1;  Mismatches    4;  Indels   43;  Gaps    3;

Qy          4 TLKAGDNLKIKQ--------------FTYSLKKDLTDLTSVGTEKLSFSANGNKVNITSD 49
              ||||||||||||              ||||||||||||||| |||||| ||||||||||
Db        106 TLKAGDNLKIKQNTNKNTNENTNDSSFTYSLKKDLTDLTSVETEKLSFGANGNKVNITSD 165

Qy         50 TKGLNFAKETAGTNGDTTVHLNGIGSTLTD----------------------RAASVKDV 87
              ||||||||||||||||||||||||||||||                      |||||||
Db        166 TKGLNFAKETAGTNGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDKKKRAASVKDV 225

Qy         88 LNAGWNIKGVK-------NVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKT 140
              ||||||||||       ||||| |||||||||||||||||||||||:|||||||||
Db        226 LNAGWNIKGVKPGTTASDNVDFVHTYDTVEFLSADTKTTTVNVESKDNGKRTEVKIGAKT 285

Qy        141 SVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADK 200
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        286 SVIKEKDGKLVTGKGKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADK 345

Qy        201 F 201
              |
Db        346 F 346


RESULT 13
AAY23739
ID   AAY23739 standard; Protein; 594 AA.
XX
AC   AAY23739;
XX
DT   08-SEP-1999  (first entry)
XX
DE   A surface protein of Neisseria meningitidis.
XX
KW   Surface protein; surface glycoprotein; infection; vaccine;
KW   immunoreactive peptide.
XX
OS   Neisseria meningitidis.
XX
PN   WO9931132-A1.
XX
PD   24-JUN-1999.
XX
PF   14-DEC-1998;   98WO-AU01031.
XX
PR   12-DEC-1997;   97GB-0026398.
XX
PA   (ISIS-) ISIS INNOVATION LTD.
PA   (UYQU ) UNIV QUEENSLAND.
XX
PI   Jennings MP,  Moxon ER,  Peak IRA;
XX
DR   WPI; 1999-418754/35.
DR   N-PSDB; AAX85791.
XX
PT   Neisseria meningitidis surface proteins useful for treating N.
PT   meningitidis infections
XX
```

```
PS   Claim 1; Page 95-97; 132pp; English.
XX
CC   The present sequence represents a surface protein of Neiserria
CC   meningitidis which is approximately 62 kDa. The N. meningitidis
CC   surface glycoproteins, nucleic acids, the primers and optionally
CC   a thermostable polymerase, or antibodies are useful in a kit for
CC   the detection or diagnosis of N. meningitidis infection in humans.
CC   The N. meningitidis surface glycoproteins can also be used to
CC   prevent or treat N. meningitidis infection in humans, especially
CC   in the form of vaccines. The proteins and antibodies can also
CC   be used to identify immunoreactive peptides.
XX
SQ   Sequence   594 AA;

  Query Match            90.7%;  Score 923.5;  DB 20;  Length 594;
  Best Local Similarity  81.2%;  Pred. No. 6.3e-70;
  Matches  194;  Conservative    0;  Mismatches    6;  Indels   39;  Gaps    3;

Qy          2 ANTLKAGDNLKIKQ----------FTYSLKKDLTDLTSVGTEKLSFSANGNKVNITSDTK 51
              | ||||||||||||          ||||||||||||||| |||||| ||||||||||||
Db        103 AITLKAGDNLKIKQNTNENTNDSSFTYSLKKDLTDLTSVETEKLSFGANGNKVNITSDTK 162

Qy         52 GLNFAKETAGTNGDTTVHLNGIGSTLTD----------------------RAASVKDVLN 89
              |||||||||||||| ||||||||||||                      ||||||||||
Db        163 GLNFAKETAGTNGDPTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLN 222

Qy         90 AGWNIKGVK-------NVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSV 142
              |||||||||       ||||||||||||||||||||||||||||||||||||||||||||
Db        223 AGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSV 282

Qy        143 IKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKF 201
              |||||||||||| | ||||||||||||||||||||||||||||||||||||||||||||
Db        283 IKEKDGKLVTGKGKDENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKF 341


RESULT 14
AAU06179
ID   AAU06179 standard; Protein; 594 AA.
XX
AC   AAU06179;
XX
DT   24-OCT-2001  (first entry)
XX
DE   N. meningitidis BZ198 surface antigen NhhA polypeptide sequence.
XX
KW   Surface antigen NhhA; meningococcal disease; meningitis vaccine.
XX
OS   Neisseria meningitidis strain BZ198.
XX
FH   Key             Location/Qualifiers
FT   Region          1..50
FT                   /label= C1
FT                   /note= "Conserved region 1"
FT   Region          51..104
FT                   /label= V1
FT                   /note= "Variable region 1"
FT   Region          105..116
FT                   /label= C2
FT                   /note= "Conserved region 2"
FT   Region          117..126
FT                   /label= V2
FT                   /note= "Variable region 2"
FT   Region          127..190
FT                   /label= C3
FT                   /note= "Conserved region 3"
FT   Region          191..212
FT                   /label= V3
FT                   /note= "Variable region 3"
FT   Region          213..231
FT                   /label= C4
FT                   /note= "Conserved region 4"
FT   Region          232..238
FT                   /label= V4
FT                   /note= "Variable region 4"
FT   Region          239..594
FT                   /label= C5
FT                   /note= "Conserved region 5"
XX
PN   WO200155182-A1.
XX
PD   02-AUG-2001.
XX
PF   25-JAN-2001; 2001WO-AU00069.
XX
PR   25-JAN-2000; 2000US-0177917.
XX
PA   (UYQU ) UNIV QUEENSLAND.
XX
PI   Peak IRA,  Jennings MP;
XX
DR   WPI; 2001-488774/53.
DR   N-PSDB; AAS09169.
XX
PT   New NhhA surface antigen polypeptides and polynucleotides from
PT   Neisseria meningitidis, useful in producing vaccines for treating or
PT   preventing broad spectrum of Neisseria meningitidis -
XX
PS   Claim 9; Fig 1; 91pp; English.
XX
CC   The present invention relates to the isolation of novel Neisseria
CC   meningitidis mutant polypeptides of the surface antigen NhhA
CC   (AAU06182-AAU06186). The modified or mutant NhhA polypeptides are
CC   characterised by deletions of non-conserved amino acids, particularly
CC   the deletion of variable regions. The deletion mutants are useful in
CC   diagnostics, therapeutic and prophylactic vaccines against a broader
CC   spectrum of N. meningitidis, and in designing and/or screening of
CC   medicaments. The mutant proteins when used as a vaccine can effectively
CC   immunise against a broader spectrum of N. meningitidis strains than
CC   would be expected from a corresponding wild-type surface antigen.
CC   The present sequence representing the wild type surface antigen NhhA
CC   from N. meningitidis strain BZ198 is 1 of 10 NhhA polypeptide sequences
CC   (AAU06171-AAU06180) from 10 different N. meningitidis strains given in
CC   the present invention.
XX
SQ   Sequence   594 AA;

  Query Match            90.7%;  Score 923.5;  DB 22;  Length 594;
  Best Local Similarity  81.2%;  Pred. No. 6.3e-70;
  Matches  194;  Conservative    0;  Mismatches    6;  Indels   39;  Gaps    3;

Qy          2 ANTLKAGDNLKIKQ----------FTYSLKKDLTDLTSVGTEKLSFSANGNKVNITSDTK 51
              | ||||||||||||          ||||||||||||||| |||||| ||||||||||||
Db        103 AITLKAGDNLKIKQNTNENTNDSSFTYSLKKDLTDLTSVETEKLSFGANGNKVNITSDTK 162

Qy         52 GLNFAKETAGTNGDTTVHLNGIGSTLTD----------------------RAASVKDVLN 89
              |||||||||||||| ||||||||||||                      ||||||||||
Db        163 GLNFAKETAGTNGDPTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLN 222

Qy         90 AGWNIKGVK-------NVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSV 142
              |||||||||       ||||||||||||||||||||||||||||||||||||||||||||
Db        223 AGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSV 282

Qy        143 IKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKF 201
              |||||||||||| | ||||||||||||||||||||||||||||||||||||||||||||
Db        283 IKEKDGKLVTGKGKDENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKF 341


RESULT 15
AAY23740
ID   AAY23740 standard; Protein; 594 AA.
XX
AC   AAY23740;
XX
DT   08-SEP-1999  (first entry)
XX
DE   A surface protein of Neisseria meningitidis.
```

XX
KW Surface protein; surface glycoprotein; infection; vaccine;
KW immunoreactive peptide.
XX
OS Neisseria meningitidis.
XX
PN WO9931132-A1.
XX
PD 24-JUN-1999.
XX
PF 14-DEC-1998; 98WO-AU01031.
XX
PR 12-DEC-1997; 97GB-0026398.
XX
PA (ISIS-) ISIS INNOVATION LTD.
PA (UYQU) UNIV QUEENSLAND.
XX
PI Jennings MP, Moxon ER, Peak IRA;
XX
DR WPI; 1999-418754/35.
DR N-PSDB; AAX85792.
XX
PT Neisseria meningitidis surface proteins useful for treating N.
PT meningitidis infections
XX
PS Claim 1; Page 100-101; 132pp; English.
XX
CC The present sequence represents a surface protein of Neiserria
CC meningitidis which is approximately 62 kDa. The N. meningitidis
CC surface glycoproteins, nucleic acids, the primers and optionally
CC a thermostable polymerase, or antibodies are useful in a kit for
CC the detection or diagnosis of N. meningitidis infection in humans.
CC The N. meningitidis surface glycoproteins can also be used to
CC prevent or treat N. meningitidis infection in humans, especially
CC in the form of vaccines. The proteins and antibodies can also
CC be used to identify immunoreactive peptides.
XX
SQ Sequence 594 AA;

Query Match 90.6%; Score 922.5; DB 20; Length 594;
Best Local Similarity 81.0%; Pred. No. 7.7e-70;
Matches 192; Conservative 3; Mismatches 3; Indels 39; Gaps 3;

```
Qy          4 TLKAGDNLKIKQ----------FTYSLKKDLTDLTSVGTEKLSFSANGNKVNITSDTKGL 53
              |||||||||||||          ||||||||||||||||||||||| ||||||||||||
Db        105 TLKAGDNLKIKQNTNENTNASSFTYSLKKDLTDLTSVGTEKLSFSANSNKVNITSDTKGL 164

Qy         54 NFAKETAGTNGDTTVHLNGIGSTLTD----------------------RAASVKDVLNAG 91
              ||||:|| |||||||||||||||||                       ||||||||||||
Db        165 NFAKKTAETNGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAG 224

Qy         92 WNIKGVK-------NVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIK 144
              |||||||       ||||||||||||||||||||||||||||||:|||||||||||||
Db        225 WNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKRTEVKIGAKTSVIK 284

Qy        145 EKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKF 201
              |||||||||||||| ||||:|||||||||||||||||||||||||||||||||||||
Db        285 EKDGKLVTGKDKGENDSSTDKGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKF 341
```

Search completed: October 6, 2003, 09:22:46
Job time : 18.261 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: October 6, 2003, 09:14:45 ; Search time 5.37514 Seconds
(without alignments)
1582.188 Million cell updates/sec

Title: US-09-771-382-39
Perfect score: 1018
Sequence: 1 SANTLKAGDNLKIKQFTYSL.....AGWRMKTTTANGQTGQADKF 201

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	957.5	94.1	591	3	US-09-377-155-11	Sequence 11, Appl
2	957.5	94.1	591	3	US-09-377-155-21	Sequence 21, Appl
3	957.5	94.1	591	4	US-09-669-974-11	Sequence 11, Appl
4	957.5	94.1	591	4	US-09-669-974-21	Sequence 21, Appl
5	953.5	93.7	592	3	US-09-377-155-2	Sequence 2, Appli
6	953.5	93.7	592	4	US-09-669-974-2	Sequence 2, Appli
7	925.5	90.9	598	3	US-09-377-155-5	Sequence 5, Appli
8	925.5	90.9	598	4	US-09-669-974-5	Sequence 5, Appli
9	925.5	90.9	599	3	US-09-377-155-15	Sequence 15, Appl
10	925.5	90.9	599	4	US-09-669-974-15	Sequence 15, Appl
11	923.5	90.7	594	3	US-09-377-155-7	Sequence 7, Appli
12	923.5	90.7	594	4	US-09-669-974-7	Sequence 7, Appli
13	922.5	90.6	594	3	US-09-377-155-9	Sequence 9, Appli
14	922.5	90.6	594	4	US-09-669-974-9	Sequence 9, Appli
15	921.5	90.5	598	3	US-09-377-155-13	Sequence 13, Appl
16	921.5	90.5	598	4	US-09-669-974-13	Sequence 13, Appl
17	918.5	90.2	589	3	US-09-377-155-19	Sequence 19, Appl
18	918.5	90.2	589	4	US-09-669-974-19	Sequence 19, Appl
19	907.5	89.1	592	3	US-09-377-155-17	Sequence 17, Appl
20	907.5	89.1	592	4	US-09-669-974-17	Sequence 17, Appl
21	622.5	61.1	2353	3	US-09-377-155-33	Sequence 33, Appl
22	622.5	61.1	2353	3	US-08-913-942-4	Sequence 4, Appli
23	622.5	61.1	2353	4	US-09-669-974-33	Sequence 33, Appl
24	622.5	61.1	2411	4	US-09-268-347-36	Sequence 36, Appl
25	618	60.7	658	1	US-08-409-995-5	Sequence 5, Appli
26	618	60.7	658	3	US-08-685-467-5	Sequence 5, Appli
27	618	60.7	658	3	US-08-913-942-5	Sequence 5, Appli
28	618	60.7	1098	1	US-08-409-995-2	Sequence 2, Appli
29	618	60.7	1098	3	US-08-685-467-2	Sequence 2, Appli
30	618	60.7	1098	3	US-09-377-155-32	Sequence 32, Appl
31	618	60.7	1098	3	US-08-913-942-2	Sequence 2, Appli
32	618	60.7	1098	4	US-09-669-974-32	Sequence 32, Appl
33	618	60.7	1098	4	US-09-268-347-44	Sequence 44, Appl
34	616.5	60.6	1912	1	US-08-409-995-4	Sequence 4, Appli
35	616.5	60.6	1912	3	US-08-685-467-4	Sequence 4, Appli
36	616.5	60.6	2354	4	US-09-268-347-47	Sequence 47, Appl
37	605.5	59.5	607	1	US-08-409-995-6	Sequence 6, Appli
38	605.5	59.5	607	3	US-08-685-467-6	Sequence 6, Appli
39	605.5	59.5	607	3	US-08-913-942-6	Sequence 6, Appli
40	555	54.5	1094	4	US-09-268-347-32	Sequence 32, Appl
41	417	41.0	679	3	US-08-913-942-15	Sequence 15, Appl
42	417	41.0	679	4	US-09-268-347-26	Sequence 26, Appl
43	327.5	32.2	1004	4	US-09-268-347-30	Sequence 30, Appl
44	320	31.4	1002	4	US-09-268-347-24	Sequence 24, Appl
45	311.5	30.6	1104	4	US-09-268-347-28	Sequence 28, Appl

ALIGNMENTS

RESULT 1
US-09-377-155-11
; Sequence 11, Application US/09377155
; Patent No. 6197312
; GENERAL INFORMATION:
; APPLICANT: PEAK, Ian Richard Anselm
; APPLICANT: JENNINGS, Michael Paul
; APPLICANT: MOXON, E. Richard
; TITLE OF INVENTION: NOVEL SURFACE ANTIGEN
; FILE REFERENCE: 065064/0128
; CURRENT APPLICATION NUMBER: US/09/377,155
; CURRENT FILING DATE: 1999-08-19
; PRIOR APPLICATION NUMBER: PCT/AU98/01031
; PRIOR FILING DATE: 1998-12-14
; PRIOR APPLICATION NUMBER: GB 9726398.2
; PRIOR FILING DATE: 1997-12-12
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 591
; TYPE: PRT
; ORGANISM: Neisseria meningitidis
US-09-377-155-11

Query Match 94.1%; Score 957.5; DB 3; Length 591;
Best Local Similarity 85.7%; Pred. No. 2.4e-84;
Matches 198; Conservative 0; Mismatches 0; Indels 33; Gaps 3;

```
Qy          4 TLKAGDNLKIKQ----FTYSLKKDLTDLTSVGTEKLSFSANGNKVNITSDTKGLNFAKET 59
              ||||||||||||    ||||||||||||||||||||||||||||||||||||||||||||
Db        109 TLKAGDNLKIKQNGTNFTYSLKKDLTDLTSVGTEKLSFSANGNKVNITSDTKGLNFAKET 168

Qy         60 AGTNGDTTVHLNGIGSTLTD----------------------RAASVKDVLNAGWNIKGV 97
              ||||||||||||||||||||                      ||||||||||||||||||
Db        169 AGTNGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGV 228

Qy         98 K-------NVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKL 150
              |       ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        229 KPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKL 288

Qy        151 VTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKF 201
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        289 VTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKF 339
```

RESULT 2
US-09-377-155-21
; Sequence 21, Application US/09377155
; Patent No. 6197312
; GENERAL INFORMATION:

```
;  APPLICANT:  PEAK, Ian Richard Anselm
;  APPLICANT:  JENNINGS, Michael Paul
;  APPLICANT:  MOXON, E. Richard
;  TITLE OF INVENTION:  NOVEL SURFACE ANTIGEN
;  FILE REFERENCE:  065064/0128
;  CURRENT APPLICATION NUMBER: US/09/377,155
;  CURRENT FILING DATE:  1999-08-19
;  PRIOR APPLICATION NUMBER:  PCT/AU98/01031
;  PRIOR FILING DATE:  1998-12-14
;  PRIOR APPLICATION NUMBER:  GB 9726398.2
;  PRIOR FILING DATE:  1997-12-12
;  NUMBER OF SEQ ID NOS:  33
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 21
;   LENGTH: 591
;   TYPE: PRT
;   ORGANISM: Neisseria meningitidis
US-09-377-155-21

  Query Match             94.1%;  Score 957.5;  DB 3;  Length 591;
  Best Local Similarity   85.7%;  Pred. No. 2.4e-84;
  Matches  198;  Conservative    0;  Mismatches    0;  Indels   33;  Gaps    3;

Qy          4 TLKAGDNLKIKQ----FTYSLKKDLTDLTSVGTEKLSFSANGNKVNITSDTKGLNFAKET 59
              ||||||||||||    ||||||||||||||||||||||||||||||||||||||||||||
Db        109 TLKAGDNLKIKQNGTNFTYSLKKDLTDLTSVGTEKLSFSANGNKVNITSDTKGLNFAKET 168

Qy         60 AGTNGDTTVHLNGIGSTLTD----------------------RAASVKDVLNAGWNIKGV 97
              ||||||||||||||||||||                      ||||||||||||||||||
Db        169 AGTNGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGV 228

Qy         98 K-------NVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKL 150
              |       ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        229 KPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKL 288

Qy        151 VTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKF 201
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        289 VTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKF 339


RESULT 3
US-09-669-974-11
; Sequence 11, Application US/09669974
; Patent No. 6333173
; GENERAL INFORMATION:
;  APPLICANT: PEAK, Ian Richard Anselm
;  APPLICANT:  JENNINGS, Michael Paul
;  APPLICANT:  MOXON, E. Richard
;  TITLE OF INVENTION: NOVEL SURFACE ANTIGEN
;  FILE REFERENCE: 065064/0128
;  CURRENT APPLICATION NUMBER: US/09/669,974
;  CURRENT FILING DATE:  2000-09-26
;  PRIOR APPLICATION NUMBER: US 09/377,155
;  PRIOR FILING DATE: 1999-08-19
;  PRIOR APPLICATION NUMBER: PCT/AU98/01031
;  PRIOR FILING DATE: 1998-12-14
;  PRIOR APPLICATION NUMBER: GB 9726398.2
;  PRIOR FILING DATE: 1997-12-12
;  NUMBER OF SEQ ID NOS: 33
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
;   LENGTH: 591
;   TYPE: PRT
;   ORGANISM: Neisseria meningitidis
US-09-669-974-11

  Query Match             94.1%;  Score 957.5;  DB 4;  Length 591;
  Best Local Similarity   85.7%;  Pred. No. 2.4e-84;
  Matches  198;  Conservative    0;  Mismatches    0;  Indels   33;  Gaps    3;

Qy          4 TLKAGDNLKIKQ----FTYSLKKDLTDLTSVGTEKLSFSANGNKVNITSDTKGLNFAKET 59
              ||||||||||||    ||||||||||||||||||||||||||||||||||||||||||||
Db        109 TLKAGDNLKIKQNGTNFTYSLKKDLTDLTSVGTEKLSFSANGNKVNITSDTKGLNFAKET 168

Qy         60 AGTNGDTTVHLNGIGSTLTD----------------------RAASVKDVLNAGWNIKGV 97
              ||||||||||||||||||||                      ||||||||||||||||||
Db        169 AGTNGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGV 228

Qy         98 K-------NVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKL 150
              |       ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        229 KPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKL 288

Qy        151 VTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKF 201
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        289 VTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKF 339


RESULT 4
US-09-669-974-21
; Sequence 21, Application US/09669974
; Patent No. 6333173
; GENERAL INFORMATION:
;  APPLICANT: PEAK, Ian Richard Anselm
;  APPLICANT:  JENNINGS, Michael Paul
;  APPLICANT:  MOXON, E. Richard
;  TITLE OF INVENTION: NOVEL SURFACE ANTIGEN
;  FILE REFERENCE: 065064/0128
;  CURRENT APPLICATION NUMBER: US/09/669,974
;  CURRENT FILING DATE:  2000-09-26
;  PRIOR APPLICATION NUMBER: US 09/377,155
;  PRIOR FILING DATE: 1999-08-19
;  PRIOR APPLICATION NUMBER: PCT/AU98/01031
;  PRIOR FILING DATE: 1998-12-14
;  PRIOR APPLICATION NUMBER: GB 9726398.2
;  PRIOR FILING DATE: 1997-12-12
;  NUMBER OF SEQ ID NOS: 33
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 21
;   LENGTH: 591
;   TYPE: PRT
;   ORGANISM: Neisseria meningitidis
US-09-669-974-21

  Query Match             94.1%;  Score 957.5;  DB 4;  Length 591;
  Best Local Similarity   85.7%;  Pred. No. 2.4e-84;
  Matches  198;  Conservative    0;  Mismatches    0;  Indels   33;  Gaps    3;

Qy          4 TLKAGDNLKIKQ----FTYSLKKDLTDLTSVGTEKLSFSANGNKVNITSDTKGLNFAKET 59
              ||||||||||||    ||||||||||||||||||||||||||||||||||||||||||||
Db        109 TLKAGDNLKIKQNGTNFTYSLKKDLTDLTSVGTEKLSFSANGNKVNITSDTKGLNFAKET 168

Qy         60 AGTNGDTTVHLNGIGSTLTD----------------------RAASVKDVLNAGWNIKGV 97
              ||||||||||||||||||||                      ||||||||||||||||||
Db        169 AGTNGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGV 228

Qy         98 K-------NVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKL 150
              |       ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        229 KPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKL 288

Qy        151 VTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKF 201
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        289 VTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKF 339


RESULT 5
US-09-377-155-2
; Sequence 2, Application US/09377155
; Patent No. 6197312
; GENERAL INFORMATION:
;  APPLICANT:  PEAK, Ian Richard Anselm
;  APPLICANT:  JENNINGS, Michael Paul
;  APPLICANT:  MOXON, E. Richard
;  TITLE OF INVENTION:  NOVEL SURFACE ANTIGEN
;  FILE REFERENCE:  065064/0128
;  CURRENT APPLICATION NUMBER: US/09/377,155
;  CURRENT FILING DATE:  1999-08-19
```

```
;  PRIOR APPLICATION NUMBER:  PCT/AU98/01031
;  PRIOR FILING DATE:  1998-12-14
;  PRIOR APPLICATION NUMBER:  GB 9726398.2
;  PRIOR FILING DATE:  1997-12-12
;  NUMBER OF SEQ ID NOS:  33
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 592
;   TYPE: PRT
;   ORGANISM: Neisseria meningitidis
US-09-377-155-2

  Query Match             93.7%;  Score 953.5;  DB 3;  Length 592;
  Best Local Similarity   85.3%;  Pred. No. 5.8e-84;
  Matches  197;  Conservative    0;  Mismatches    1;  Indels   33;  Gaps    3;

Qy          4 TLKAGDNLKIKQ----FTYSLKKDLTDLTSVGTEKLSFSANGNKVNITSDTKGLNFAKET 59
              ||||||||||||    ||||||||||||||||||||||||||||||||||||||||||||
Db        110 TLKAGDNLKIKQNGTNFTYSLKKDLTDLTSVGTEKLSFSANGNKVNITSDTKGLNFAKET 169

Qy         60 AGTNGDTTVHLNGIGSTLTD----------------------RAASVKDVLNAGWNIKGV 97
              ||||||||||||||||||||                      ||||||||||||||||||
Db        170 AGTNGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGV 229

Qy         98 K--------NVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKL 150
              |        ||||||||||||||||||||||||||||||||||||| ||||||||||||
Db        230 KPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGVKTSVIKEKDGKL 289

Qy        151 VTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKF 201
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        290 VTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKF 340


RESULT 6
US-09-669-974-2
; Sequence 2, Application US/09669974
; Patent No. 6333173
; GENERAL INFORMATION:
;  APPLICANT: PEAK, Ian Richard Anselm
;  APPLICANT:  JENNINGS, Michael Paul
;  APPLICANT:  MOXON, E. Richard
;  TITLE OF INVENTION: NOVEL SURFACE ANTIGEN
;  FILE REFERENCE: 065064/0128
;  CURRENT APPLICATION NUMBER: US/09/669,974
;  CURRENT FILING DATE:  2000-09-26
;  PRIOR APPLICATION NUMBER: US 09/377,155
;  PRIOR FILING DATE: 1999-08-19
;  PRIOR APPLICATION NUMBER: PCT/AU98/01031
;  PRIOR FILING DATE: 1998-12-14
;  PRIOR APPLICATION NUMBER: GB 9726398.2
;  PRIOR FILING DATE: 1997-12-12
;  NUMBER OF SEQ ID NOS: 33
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 592
;   TYPE: PRT
;   ORGANISM: Neisseria meningitidis
US-09-669-974-2

  Query Match             93.7%;  Score 953.5;  DB 4;  Length 592;
  Best Local Similarity   85.3%;  Pred. No. 5.8e-84;
  Matches  197;  Conservative    0;  Mismatches    1;  Indels   33;  Gaps    3;

Qy          4 TLKAGDNLKIKQ----FTYSLKKDLTDLTSVGTEKLSFSANGNKVNITSDTKGLNFAKET 59
              ||||||||||||    ||||||||||||||||||||||||||||||||||||||||||||
Db        110 TLKAGDNLKIKQNGTNFTYSLKKDLTDLTSVGTEKLSFSANGNKVNITSDTKGLNFAKET 169

Qy         60 AGTNGDTTVHLNGIGSTLTD----------------------RAASVKDVLNAGWNIKGV 97
              ||||||||||||||||||||                      ||||||||||||||||||
Db        170 AGTNGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGV 229

Qy         98 K-------NVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKL 150
              |       ||||||||||||||||||||||||||||||||||||| ||||||||||||
Db        230 KPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGVKTSVIKEKDGKL 289

Qy        151 VTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKF 201
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        290 VTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKF 340


RESULT 7
US-09-377-155-5
; Sequence 5, Application US/09377155
; Patent No. 6197312
; GENERAL INFORMATION:
;  APPLICANT:  PEAK, Ian Richard Anselm
;  APPLICANT:  JENNINGS, Michael Paul
;  APPLICANT:  MOXON, E. Richard
;  TITLE OF INVENTION:  NOVEL SURFACE ANTIGEN
;  FILE REFERENCE:  065064/0128
;  CURRENT APPLICATION NUMBER: US/09/377,155
;  CURRENT FILING DATE:  1999-08-19
;  PRIOR APPLICATION NUMBER:  PCT/AU98/01031
;  PRIOR FILING DATE:  1998-12-14
;  PRIOR APPLICATION NUMBER:  GB 9726398.2
;  PRIOR FILING DATE:  1997-12-12
;  NUMBER OF SEQ ID NOS:  33
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 5
;   LENGTH: 598
;   TYPE: PRT
;   ORGANISM: Neisseria meningitidis
US-09-377-155-5

  Query Match             90.9%;  Score 925.5;  DB 3;  Length 598;
  Best Local Similarity   79.8%;  Pred. No. 3e-81;
  Matches  194;  Conservative    1;  Mismatches    5;  Indels   43;  Gaps    3;

Qy          2 ANTLKAGDNLKIKQ--------------FTYSLKKDLTDLTSVGTEKLSFSANGNKVNIT 47
              | ||||||||||||               ||||||||||||||| |||||| |||||||||
Db        103 AITLKAGDNLKIKQNTNENTNENTNDSSFTYSLKKDLTDLTSVETEKLSFGANGNKVNIT 162

Qy         48 SDTKGLNFAKETAGTNGDTTVHLNGIGSTLTD----------------------RAASVK 85
              ||||||||||||||||| ||||||||||||                       ||||||
Db        163 SDTKGLNFAKETAGTNGDPTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVK 222

Qy         86 DVLNAGWNIKGVK-------NVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGA 138
              ||||||||||||       |||||||||||||||||||||||||||||:|||||||
Db        223 DVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKRTEVKIGA 282

Qy        139 KTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQA 198
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        283 KTSVIKEKDGKLVTGKGKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQA 342

Qy        199 DKF 201
              |||
Db        343 DKF 345


RESULT 8
US-09-669-974-5
; Sequence 5, Application US/09669974
; Patent No. 6333173
; GENERAL INFORMATION:
;  APPLICANT: PEAK, Ian Richard Anselm
;  APPLICANT:  JENNINGS, Michael Paul
;  APPLICANT:  MOXON, E. Richard
;  TITLE OF INVENTION: NOVEL SURFACE ANTIGEN
;  FILE REFERENCE: 065064/0128
;  CURRENT APPLICATION NUMBER: US/09/669,974
;  CURRENT FILING DATE:  2000-09-26
;  PRIOR APPLICATION NUMBER: US 09/377,155
;  PRIOR FILING DATE: 1999-08-19
;  PRIOR APPLICATION NUMBER: PCT/AU98/01031
;  PRIOR FILING DATE: 1998-12-14
;  PRIOR APPLICATION NUMBER: GB 9726398.2
```

```
;  PRIOR FILING DATE: 1997-12-12
;  NUMBER OF SEQ ID NOS: 33
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
;   LENGTH: 598
;   TYPE: PRT
;   ORGANISM: Neisseria meningitidis
US-09-669-974-5

  Query Match             90.9%;  Score 925.5;  DB 4;  Length 598;
  Best Local Similarity   79.8%;  Pred. No. 3e-81;
  Matches  194;  Conservative    1;  Mismatches    5;  Indels   43;  Gaps    3;

Qy          2 ANTLKAGDNLKIKQ--------------FTYSLKKDLTDLTSVGTEKLSFSANGNKVNIT 47
              | ||||||||||||              |||||||||||||| |||||| |||||||||
Db        103 AITLKAGDNLKIKQNTNENTNENTNDSSFTYSLKKDLTDLTSVETEKLSFGANGNKVNIT 162

Qy         48 SDTKGLNFAKETAGTNGDTTVHLNGIGSTLTD----------------------RAASVK 85
              |||||||||||||||| ||||||||||||                        ||||||
Db        163 SDTKGLNFAKETAGTNGDPTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVK 222

Qy         86 DVLNAGWNIKGVK-------NVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGA 138
              ||||||||||||       |||||||||||||||||||||||||||||:|||||||
Db        223 DVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKRTEVKIGA 282

Qy        139 KTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQA 198
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        283 KTSVIKEKDGKLVTGKGKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQA 342

Qy        199 DKF 201
              |||
Db        343 DKF 345


RESULT 9
US-09-377-155-15
; Sequence 15, Application US/09377155
; Patent No. 6197312
; GENERAL INFORMATION:
;  APPLICANT:  PEAK, Ian Richard Anselm
;  APPLICANT:  JENNINGS, Michael Paul
;  APPLICANT:  MOXON, E. Richard
;  TITLE OF INVENTION:  NOVEL SURFACE ANTIGEN
;  FILE REFERENCE:  065064/0128
;  CURRENT APPLICATION NUMBER: US/09/377,155
;  CURRENT FILING DATE:  1999-08-19
;  PRIOR APPLICATION NUMBER:  PCT/AU98/01031
;  PRIOR FILING DATE:  1998-12-14
;  PRIOR APPLICATION NUMBER:  GB 9726398.2
;  PRIOR FILING DATE:  1997-12-12
;  NUMBER OF SEQ ID NOS:  33
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 15
;   LENGTH: 599
;   TYPE: PRT
;   ORGANISM: Neisseria meningitidis
US-09-377-155-15

  Query Match             90.9%;  Score 925.5;  DB 3;  Length 599;
  Best Local Similarity   80.1%;  Pred. No. 3e-81;
  Matches  193;  Conservative    1;  Mismatches    4;  Indels   43;  Gaps    3;

Qy          4 TLKAGDNLKIKQ--------------FTYSLKKDLTDLTSVGTEKLSFSANGNKVNITSD 49
              ||||||||||||              |||||||||||||| |||||| ||||||||||
Db        106 TLKAGDNLKIKQNTNKNTNENTNDSSFTYSLKKDLTDLTSVETEKLSFGANGNKVNITSD 165

Qy         50 TKGLNFAKETAGTNGDTTVHLNGIGSTLTD----------------------RAASVKDV 87
              ||||||||||||||||||||||||||||                        ||||||||
Db        166 TKGLNFAKETAGTNGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDKKKRAASVKDV 225

Qy         88 LNAGWNIKGVK-------NVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKT 140
              ||||||||||||       ||||| |||||||||||||||||||||||:|||||||||
Db        226 LNAGWNIKGVKPGTTASDNVDFVHTYDTVEFLSADTKTTTVNVESKDNGKRTEVKIGAKT 285

Qy        141 SVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADK 200
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        286 SVIKEKDGKLVTGKGKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADK 345

Qy        201 F 201
              |
Db        346 F 346


RESULT 10
US-09-669-974-15
; Sequence 15, Application US/09669974
; Patent No. 6333173
; GENERAL INFORMATION:
;  APPLICANT: PEAK, Ian Richard Anselm
;  APPLICANT:  JENNINGS, Michael Paul
;  APPLICANT:  MOXON, E. Richard
;  TITLE OF INVENTION: NOVEL SURFACE ANTIGEN
;  FILE REFERENCE: 065064/0128
;  CURRENT APPLICATION NUMBER: US/09/669,974
;  CURRENT FILING DATE:  2000-09-26
;  PRIOR APPLICATION NUMBER: US 09/377,155
;  PRIOR FILING DATE: 1999-08-19
;  PRIOR APPLICATION NUMBER: PCT/AU98/01031
;  PRIOR FILING DATE: 1998-12-14
;  PRIOR APPLICATION NUMBER: GB 9726398.2
;  PRIOR FILING DATE: 1997-12-12
;  NUMBER OF SEQ ID NOS: 33
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
;   LENGTH: 599
;   TYPE: PRT
;   ORGANISM: Neisseria meningitidis
US-09-669-974-15

  Query Match             90.9%;  Score 925.5;  DB 4;  Length 599;
  Best Local Similarity   80.1%;  Pred. No. 3e-81;
  Matches  193;  Conservative    1;  Mismatches    4;  Indels   43;  Gaps    3;

Qy          4 TLKAGDNLKIKQ--------------FTYSLKKDLTDLTSVGTEKLSFSANGNKVNITSD 49
              ||||||||||||              |||||||||||||| |||||| ||||||||||
Db        106 TLKAGDNLKIKQNTNKNTNENTNDSSFTYSLKKDLTDLTSVETEKLSFGANGNKVNITSD 165

Qy         50 TKGLNFAKETAGTNGDTTVHLNGIGSTLTD----------------------RAASVKDV 87
              ||||||||||||||||||||||||||||                        ||||||||
Db        166 TKGLNFAKETAGTNGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDKKKRAASVKDV 225

Qy         88 LNAGWNIKGVK-------NVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKT 140
              ||||||||||||       ||||| |||||||||||||||||||||||:|||||||||
Db        226 LNAGWNIKGVKPGTTASDNVDFVHTYDTVEFLSADTKTTTVNVESKDNGKRTEVKIGAKT 285

Qy        141 SVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADK 200
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        286 SVIKEKDGKLVTGKGKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADK 345

Qy        201 F 201
              |
Db        346 F 346


RESULT 11
US-09-377-155-7
; Sequence 7, Application US/09377155
; Patent No. 6197312
; GENERAL INFORMATION:
;  APPLICANT:  PEAK, Ian Richard Anselm
;  APPLICANT:  JENNINGS, Michael Paul
;  APPLICANT:  MOXON, E. Richard
;  TITLE OF INVENTION:  NOVEL SURFACE ANTIGEN
;  FILE REFERENCE:  065064/0128
;  CURRENT APPLICATION NUMBER: US/09/377,155
;  CURRENT FILING DATE:  1999-08-19
```

```
;  PRIOR APPLICATION NUMBER:  PCT/AU98/01031
;  PRIOR FILING DATE:  1998-12-14
;  PRIOR APPLICATION NUMBER:  GB 9726398.2
;  PRIOR FILING DATE:  1997-12-12
;  NUMBER OF SEQ ID NOS:  33
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 7
;   LENGTH: 594
;   TYPE: PRT
;   ORGANISM: Neisseria meningitidis
US-09-377-155-7

  Query Match             90.7%;  Score 923.5;  DB 3;  Length 594;
  Best Local Similarity   81.2%;  Pred. No. 4.6e-81;
  Matches  194;  Conservative    0;  Mismatches    6;  Indels   39;  Gaps    3;

Qy          2 ANTLKAGDNLKIKQ----------FTYSLKKDLTDLTSVGTEKLSFSANGNKVNITSDTK 51
              | ||||||||||||          ||||||||||||||| |||||| ||||||||||||
Db        103 AITLKAGDNLKIKQNTNENTNDSSFTYSLKKDLTDLTSVETEKLSFGANGNKVNITSDTK 162

Qy         52 GLNFAKETAGTNGDTTVHLNGIGSTLTD----------------------RAASVKDVLN 89
              |||||||||||||| ||||||||||||                      ||||||||||
Db        163 GLNFAKETAGTNGDPTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLN 222

Qy         90 AGWNIKGVK-------NVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSV 142
              |||||||||       ||||||||||||||||||||||||||||||||||||||||||||
Db        223 AGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSV 282

Qy        143 IKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKF 201
              |||||||||||| | ||||||||||||||||||||||||||||||||||||||||||||
Db        283 IKEKDGKLVTGKGKDENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKF 341


RESULT 12
US-09-669-974-7
; Sequence 7, Application US/09669974
; Patent No. 6333173
; GENERAL INFORMATION:
;  APPLICANT: PEAK, Ian Richard Anselm
;  APPLICANT:  JENNINGS, Michael Paul
;  APPLICANT:  MOXON, E. Richard
;  TITLE OF INVENTION: NOVEL SURFACE ANTIGEN
;  FILE REFERENCE: 065064/0128
;  CURRENT APPLICATION NUMBER: US/09/669,974
;  CURRENT FILING DATE:  2000-09-26
;  PRIOR APPLICATION NUMBER: US 09/377,155
;  PRIOR FILING DATE: 1999-08-19
;  PRIOR APPLICATION NUMBER: PCT/AU98/01031
;  PRIOR FILING DATE: 1998-12-14
;  PRIOR APPLICATION NUMBER: GB 9726398.2
;  PRIOR FILING DATE: 1997-12-12
;  NUMBER OF SEQ ID NOS: 33
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
;   LENGTH: 594
;   TYPE: PRT
;   ORGANISM: Neisseria meningitidis
US-09-669-974-7

  Query Match             90.7%;  Score 923.5;  DB 4;  Length 594;
  Best Local Similarity   81.2%;  Pred. No. 4.6e-81;
  Matches  194;  Conservative    0;  Mismatches    6;  Indels   39;  Gaps    3;

Qy          2 ANTLKAGDNLKIKQ----------FTYSLKKDLTDLTSVGTEKLSFSANGNKVNITSDTK 51
              | ||||||||||||          ||||||||||||||| |||||| ||||||||||||
Db        103 AITLKAGDNLKIKQNTNENTNDSSFTYSLKKDLTDLTSVETEKLSFGANGNKVNITSDTK 162

Qy         52 GLNFAKETAGTNGDTTVHLNGIGSTLTD----------------------RAASVKDVLN 89
              |||||||||||||| ||||||||||||                      ||||||||||
Db        163 GLNFAKETAGTNGDPTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLN 222

Qy         90 AGWNIKGVK-------NVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSV 142
              |||||||||       ||||||||||||||||||||||||||||||||||||||||||||
Db        223 AGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSV 282

Qy        143 IKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKF 201
              |||||||||||| | ||||||||||||||||||||||||||||||||||||||||||||
Db        283 IKEKDGKLVTGKGKDENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKF 341


RESULT 13
US-09-377-155-9
; Sequence 9, Application US/09377155
; Patent No. 6197312
; GENERAL INFORMATION:
;  APPLICANT:  PEAK, Ian Richard Anselm
;  APPLICANT:  JENNINGS, Michael Paul
;  APPLICANT:  MOXON, E. Richard
;  TITLE OF INVENTION:  NOVEL SURFACE ANTIGEN
;  FILE REFERENCE:  065064/0128
;  CURRENT APPLICATION NUMBER: US/09/377,155
;  CURRENT FILING DATE:  1999-08-19
;  PRIOR APPLICATION NUMBER:  PCT/AU98/01031
;  PRIOR FILING DATE:  1998-12-14
;  PRIOR APPLICATION NUMBER:  GB 9726398.2
;  PRIOR FILING DATE:  1997-12-12
;  NUMBER OF SEQ ID NOS:  33
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 9
;   LENGTH: 594
;   TYPE: PRT
;   ORGANISM: Neisseria meningitidis
US-09-377-155-9

  Query Match             90.6%;  Score 922.5;  DB 3;  Length 594;
  Best Local Similarity   81.0%;  Pred. No. 5.8e-81;
  Matches  192;  Conservative    3;  Mismatches    3;  Indels   39;  Gaps    3;

Qy          4 TLKAGDNLKIKQ----------FTYSLKKDLTDLTSVGTEKLSFSANGNKVNITSDTKGL 53
              |||||||||||          |||||||||||||||||||||||| ||||||||||||
Db        105 TLKAGDNLKIKQNTNENTNASSFTYSLKKDLTDLTSVGTEKLSFSANSNKVNITSDTKGL 164

Qy         54 NFAKETAGTNGDTTVHLNGIGSTLTD----------------------RAASVKDVLNAG 91
              ||||:|| ||||||||||||||||                      ||||||||||||
Db        165 NFAKKTAETNGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAG 224

Qy         92 WNIKGVK-------NVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIK 144
              |||||||       |||||||||||||||||||||||||||||:||||||||||||||
Db        225 WNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKRTEVKIGAKTSVIK 284

Qy        145 EKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKF 201
              |||||||||||||| ||||:|||||||||||||||||||||||||||||||||||||
Db        285 EKDGKLVTGKDKGENDSSTDKGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKF 341


RESULT 14
US-09-669-974-9
; Sequence 9, Application US/09669974
; Patent No. 6333173
; GENERAL INFORMATION:
;  APPLICANT: PEAK, Ian Richard Anselm
;  APPLICANT:  JENNINGS, Michael Paul
;  APPLICANT:  MOXON, E. Richard
;  TITLE OF INVENTION: NOVEL SURFACE ANTIGEN
;  FILE REFERENCE: 065064/0128
;  CURRENT APPLICATION NUMBER: US/09/669,974
;  CURRENT FILING DATE:  2000-09-26
;  PRIOR APPLICATION NUMBER: US 09/377,155
;  PRIOR FILING DATE: 1999-08-19
;  PRIOR APPLICATION NUMBER: PCT/AU98/01031
;  PRIOR FILING DATE: 1998-12-14
;  PRIOR APPLICATION NUMBER: GB 9726398.2
;  PRIOR FILING DATE: 1997-12-12
;  NUMBER OF SEQ ID NOS: 33
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
```

```
;   LENGTH: 594
;   TYPE: PRT
;   ORGANISM: Neisseria meningitidis
US-09-669-974-9

  Query Match             90.6%;  Score 922.5;  DB 4;  Length 594;
  Best Local Similarity   81.0%;  Pred. No. 5.8e-81;
  Matches  192;  Conservative    3;  Mismatches    3;  Indels   39;  Gaps    3;

Qy          4 TLKAGDNLKIKQ----------FTYSLKKDLTDLTSVGTEKLSFSANGNKVNITSDTKGL 53
              ||||||||||||          |||||||||||||||||||||||| ||||||||||||
Db        105 TLKAGDNLKIKQNTNENTNASSFTYSLKKDLTDLTSVGTEKLSFSANSNKVNITSDTKGL 164

Qy         54 NFAKETAGTNGDTTVHLNGIGSTLTD----------------------RAASVKDVLNAG 91
              ||||:|| |||||||||||||||||                      ||||||||||||
Db        165 NFAKKTAETNGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAG 224

Qy         92 WNIKGVK-------NVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIK 144
              |||||||       ||||||||||||||||||||||||||||:||||||||||||||
Db        225 WNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKRTEVKIGAKTSVIK 284

Qy        145 EKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKF 201
              |||||||||||||| ||||:|||||||||||||||||||||||||||||||||||||
Db        285 EKDGKLVTGKDKGENDSSTDKGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKF 341


RESULT 15
US-09-377-155-13
; Sequence 13, Application US/09377155
; Patent No. 6197312
; GENERAL INFORMATION:
;  APPLICANT:  PEAK, Ian Richard Anselm
;  APPLICANT:  JENNINGS, Michael Paul
;  APPLICANT:  MOXON, E. Richard
;  TITLE OF INVENTION:  NOVEL SURFACE ANTIGEN
;  FILE REFERENCE:  065064/0128
;  CURRENT APPLICATION NUMBER: US/09/377,155
;  CURRENT FILING DATE:  1999-08-19
;  PRIOR APPLICATION NUMBER:  PCT/AU98/01031
;  PRIOR FILING DATE:  1998-12-14
;  PRIOR APPLICATION NUMBER:  GB 9726398.2
;  PRIOR FILING DATE:  1997-12-12
;  NUMBER OF SEQ ID NOS:  33
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 13
;   LENGTH: 598
;   TYPE: PRT
;   ORGANISM: Neisseria meningitidis
US-09-377-155-13

  Query Match             90.5%;  Score 921.5;  DB 3;  Length 598;
  Best Local Similarity   79.8%;  Pred. No. 7.3e-81;
  Matches  194;  Conservative    0;  Mismatches    6;  Indels   43;  Gaps    3;

Qy          2 ANTLKAGDNLKIKQ--------------FTYSLKKDLTDLTSVGTEKLSFSANGNKVNIT 47
              | ||||||||||||              ||||||||||||||| |||||| |||||||||
Db        103 AITLKAGDNLKIKQNTNENTNENTNDSSFTYSLKKDLTDLTSVETEKLSFGANGNKVNIT 162

Qy         48 SDTKGLNFAKETAGTNGDTTVHLNGIGSTLTD----------------------RAASVK 85
              |||||||||||||||| |||||||||||||                       ||||||
Db        163 SDTKGLNFAKETAGTNGDPTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVK 222

Qy         86 DVLNAGWNIKGVK-------NVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGA 138
              ||||||||||||        ||||||||||||||||||||||||||||||||||||||||
Db        223 DVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGA 282

Qy        139 KTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQA 198
              |||||||||||||||  | ||||||||||||||||||||||||||||||||||||||||
Db        283 KTSVIKEKDGKLVTGKGKDENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQA 342

Qy        199 DKF 201
              |||
Db        343 DKF 345
```

```
Search completed: October  6, 2003, 09:36:01
Job time : 6.37514 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: October 6, 2003, 09:13:35 ; Search time 5.75367 Seconds
(without alignments)
3359.577 Million cell updates/sec

Title: US-09-771-382-39
Perfect score: 1018
Sequence: 1 SANTLKAGDNLKIKQFTYSL.....AGWRMKTTTANGQTGQADKF 201

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_76:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	957.5	94.1	591	2	G81133	adhesin NMB0992 [i
2	903.5	88.8	592	2	A81888	probable surface f
3	374	36.7	298	2	I64138	adhesin homolog HI
4	116.5	11.4	507	2	S33192	phase-1 flagellin
5	116.5	11.4	508	2	A53465	phase 1 flagellin
6	113	11.1	585	2	F90961	flagellin [importe
7	113	11.1	585	2	F85809	hypothetical prote
8	112.5	11.1	462	2	S61332	IgA-specific metal
9	112	11.0	1536	2	A43855	high-molecular-wei
10	109.5	10.8	584	2	C48658	flagellin - Escher
11	109.5	10.8	920	2	I40614	surface array prot
12	108.5	10.7	569	2	T39577	hypothetical prote
13	105	10.3	621	2	H71713	glucose inhibited
14	105	10.3	1109	2	A56143	surface-array prot
15	105	10.3	1477	2	B43855	high-molecular-wei
16	104.5	10.3	462	2	AH1184	wall associated pr
17	104.5	10.3	733	2	T23507	hypothetical prote
18	104.5	10.3	1635	2	AI0452	hemolysin [importe
19	104.5	10.3	2044	2	AB1180	probable peptidogl
20	103	10.1	504	2	S33194	phase-1 flagellin
21	102	10.0	504	2	S33186	phase-1 flagellin
22	102	10.0	595	2	A48658	flagellin - Escher
23	102	10.0	1020	2	B86414	hypothetical prote
24	101.5	10.0	653	2	T03319	gene l12 protein -
25	100.5	9.9	550	2	S44980	flagellin - Shigel
26	100.5	9.9	2551	2	B98047	hypothetical prote
27	100	9.8	441	2	A97000	pectate lyase rela
28	99.5	9.8	1386	2	AC1533	surface protein (L
29	99	9.7	796	2	A90541	hypothetical prote
30	99	9.7	918	2	G88545	protein F59B2.12 [
31	99	9.7	926	2	AE1130	conserved hypothet
32	99	9.7	1005	2	C71513	hypothetical prote
33	98.5	9.7	798	2	S62791	probable lipoprote
34	98.5	9.7	936	2	C86546	polymorphic outer
35	98.5	9.7	936	2	B81591	polymorphic membra
36	98.5	9.7	936	2	C72078	polymorphic outer
37	98	9.6	444	2	A36924	S-layer protein pr
38	98	9.6	595	2	B48658	flagellin - Escher
39	97	9.5	353	2	T09665	peroxidase (EC 1.1
40	97	9.5	504	2	S33191	phase-1 flagellin
41	97	9.5	1361	2	T03415	S-layer protein -
42	96.5	9.5	2334	2	S32920	cell wall-associat
43	96	9.4	504	2	S33190	phase-1 flagellin
44	96	9.4	940	2	AB1744	internalin protein
45	95.5	9.4	918	2	AB2445	hypothetical prote

ALIGNMENTS

RESULT 1
G81133
adhesin NMB0992 [imported] - Neisseria meningitidis (strain MC58 serogroup B)
C;Species: Neisseria meningitidis
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001
C;Accession: G81133
R;Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B. ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Masignani, V.; Pizza, M.
Science 287, 1809-1815, 2000
A;Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.;
A;Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A;Reference number: A81000; MUID:20175755; PMID:10710307
A;Accession: G81133
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-591 <TET>
A;Cross-references: GB:AE002450; GB:AE002098; NID:g7226229; PIDN:AAF41395.1; PID:g722
A;Experimental source: serogroup B, strain MC58
C;Genetics:
A;Gene: NMB0992

Query Match 94.1%; Score 957.5; DB 2; Length 591;
Best Local Similarity 85.7%; Pred. No. 6.4e-60;
Matches 198; Conservative 0; Mismatches 0; Indels 33; Gaps 3;

```
Qy          4 TLKAGDNLKIKQ----FTYSLKKDLTDLTSVGTEKLSFSANGNKVNITSDTKGLNFAKET 59
              ||||||||||||    ||||||||||||||||||||||||||||||||||||||||||||
Db        109 TLKAGDNLKIKQNGTNFTYSLKKDLTDLTSVGTEKLSFSANGNKVNITSDTKGLNFAKET 168

Qy         60 AGTNGDTTVHLNGIGSTLTD----------------------RAASVKDVLNAGWNIKGV 97
              ||||||||||||||||||||                      ||||||||||||||||||
Db        169 AGTNGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGV 228

Qy         98 K-------NVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKL 150
              |       ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        229 KPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKL 288

Qy        151 VTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKF 201
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        289 VTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKF 339
```

RESULT 2
A81888
probable surface fibril protein NMA1200 [imported] - Neisseria meningitidis (strain Z
C;Species: Neisseria meningitidis
C;Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001
C;Accession: A81888
R;Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Mo ; Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandre
Nature 404, 502-506, 2000

```
A;Title: Complete DNA sequence of a serogroup A strain of Neisseria menigitidis Z2491.
A;Reference number: A81775; MUID:20222556; PMID:10761919
A;Accession: A81888
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-592 <PAR>
A;Cross-references: GB:AL162755; GB:AL157959; NID:g7379742; PIDN:CAB84461.1; PID:g737989
A;Experimental source: serogroup A, strain Z2491
C;Genetics:
A;Gene: NMA1200

  Query Match            88.8%;  Score 903.5;  DB 2;  Length 592;
  Best Local Similarity  79.3%;  Pred. No. 4e-56;
  Matches  188;  Conservative    3;  Mismatches    7;  Indels   39;  Gaps    3;

Qy          4 TLKAGDNLKIKQ----------FTYSLKKDLTDLTSVGTEKLSFSANGNKVNITSDTKGL 53
              ||||||||||||          |||||||||| | :| |||||| ||| |||| ||||||
Db        103 TLKAGDNLKIKQNTNENTNASSFTYSLKKDLTGLINVETEKLSFGANGKKVNIISDTKGL 162

Qy         54 NFAKETAGTNGDTTVHLNGIGSTLTD--------------------RAASVKDVLNAGWN 93
              |||||||||||||||||||||||||                    ||||:|||||||||
Db        163 NFAKETAGTNGDTTVHLNGIGSTLTDTLAGSSASHVDAGNQSTHYTRAASIKDVLNAGWN 222

Qy         94 IKGVK---------NVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIK 144
              |||||          ||||||||||||||||||||||||||||||:||||||||||||
Db        223 IKGVKTGSTTGQSENVDFVRTYDTVEFLSADTKTTTVNVESKDNGKRTEVKIGAKTSVIK 282

Qy        145 EKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKF 201
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        283 EKDGKLVTGKGKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKF 339


RESULT 3
I64138
adhesin homolog HI1732 - Haemophilus influenzae (strain Rd KW20)
C;Species: Haemophilus influenzae
C;Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 24-Oct-1997
C;Accession: I64138
R;Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A
; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J
, D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M.
Science 269, 496-512, 1995
A;Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,
A;Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
A;Reference number: A64000; MUID:95350630; PMID:7542800
A;Accession: I64138
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-298 <TIGR>
A;Cross-references: GB:U32846; GB:L42023; NID:g1574588; PID:g1574589; TIGR:HI1732

  Query Match            36.7%;  Score 374;  DB 2;  Length 298;
  Best Local Similarity  44.4%;  Pred. No. 3e-19;
  Matches   92;  Conservative   21;  Mismatches   48;  Indels   46;  Gaps    6;

Qy          4 TLKAGDNLKI------KQFTYSLKKDLTDLTSVGTEKLSFSAN-------GNKVNITSDT 50
              ||||| |||       |  |::| |||   |:  :: |:   |          ||:|||
Db         88 TLKAGKNLKAKLDQGGKSVTFALAKDLDVKTAKVSDTLTIGGNTPAAGGATPKVSITSTA 147

Qy         51 KGLNFAKETAGTNGDTTVHLNGIGSTLTD--------------------RAASVKDVLN 89
                ||  ||   ||||||  |||||: ||| |                   |||::|||||
Db        148 DGLKLAK---GTNGDTAVHLNGLASTLPDVTTNTGASTSVTFSPSDIEKTRAATIKDVLN 204

Qy         90 AGWNIKGVK-------NVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSV 142
              ||||||| |       ||| |  || |||:: |  |  | : :|:||| ||||   ||||
Db        205 AGWNIKGAKVAGGNTENVDLVAGYDNVEFITGDKNTLDVVLTAKENGKTTEVKFTPKTSV 264

Qy        143 IKEKDGKLVTGKD--KGENGSSTDEGE 167
              ||: :|||:|||       |::|:  |
Db        265 IKDNNGKLLTGKQLKDANTGTATNATE 291


RESULT 4
S33192
phase-1 flagellin - Salmonella oranienburg (fragment)
C;Species: Salmonella oranienburg
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 26-Aug-1999
C;Accession: S33192
R;Masten, B.J.; Joys, T.M.
submitted to the EMBL Data Library, September 1992
A;Description: Molecular analysis of the flagellar antigen complex of Salmonella.
A;Reference number: S33185
A;Accession: S33192
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-507 <MAS>
A;Cross-references: EMBL:Z15070; NID:g297006; PIDN:CAA78779.1; PID:g297007
C;Superfamily: flagellin

  Query Match            11.4%;  Score 116.5;  DB 2;  Length 507;
  Best Local Similarity  31.1%;  Pred. No. 0.69;
  Matches   59;  Conservative   27;  Mismatches   73;  Indels   31;  Gaps   11;

Qy          4 TLKAGDNLKIKQFTYSLKKDLTDLTSVGTEKLSFSANGNKVNITSDTKGLNFAKET-AGT 62
              |  | :|  :  |  :     ||      : :: |  | ||  | | ||::|  :| ||
Db        238 TADAQNNTAVDLFKSTKSAAGTD----DAKAIATSIKGGKVGDTFDYKGVSFTIDTKAGD 293

Qy         63 NGDTTVH--LNG--IGSTLTDRAASVKDVLNAGWNIKGVKNVDFVRTYDTV---EFLSAD 115
              :|: ||   :||  :  |::|  ||  || :|   |:  |:|     | :|   :|  ||
Db        294 DGNGTVSTTINGEKVTLTISDIGASATDVNSA--KIQSSKDV-----YTSVVSGQFTFAD 346

Qy        116 TKTTTVNVESKD----NGKKTEVKI---GAKTSVIKEKDGKLVTGK----DKGENGSSTD 164
               ||   : :  |    |  | | ||   ||: :     |   : ||    ||  :| ||
Db        347 -KTKNESAKLSDLEANNAVKGESKITVNGAEYTANAAGDKVTLAGKTMFIDKTASGVSTL 405

Qy        165 EGEGLVTAKE 174
              |    ||:
Db        406 INEDAAAAKK 415


RESULT 5
A53465
phase 1 flagellin - Salmonella banana
C;Species: Salmonella banana
C;Date: 04-Sep-1997 #sequence_revision 04-Sep-1997 #text_change 26-Aug-1999
C;Accession: A53465
R;Li, J.; Nelson, K.; McWhorter, A.C.; Whittam, T.S.; Selander, R.K.
Proc. Natl. Acad. Sci. U.S.A. 91, 2552-2556, 1994
A;Title: Recombinational basis of serovar diversity in Salmonella enterica.
A;Reference number: A53465; MUID:94195780; PMID:8146152
A;Accession: A53465
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-508 <RES>
A;Cross-references: EMBL:U06202; NID:g476243; PIDN:AAA17861.1; PID:g476244
C;Genetics:
A;Gene: fliC
C;Superfamily: flagellin

  Query Match            11.4%;  Score 116.5;  DB 2;  Length 508;
  Best Local Similarity  31.1%;  Pred. No. 0.69;
  Matches   59;  Conservative   27;  Mismatches   73;  Indels   31;  Gaps   11;

Qy          4 TLKAGDNLKIKQFTYSLKKDLTDLTSVGTEKLSFSANGNKVNITSDTKGLNFAKET-AGT 62
              |  | :|  :  |  :     ||      : :: |  | ||  | | ||::|  :| ||
Db        239 TADAQNNTAVDLFKSTKSAAGTD----DAKAIATSIKGGKVGDTFDYKGVSFTIDTKAGD 294

Qy         63 NGDTTVH--LNG--IGSTLTDRAASVKDVLNAGWNIKGVKNVDFVRTYDTV---EFLSAD 115
              :|: ||   :||  :  |::|  ||  || :|   |:  |:|     | :|   :|  ||
Db        295 DGNGTVSTTINGEKVTLTISDIGASATDVNSA--KIQSSKDV-----YTSVVSGQFTFAD 347

Qy        116 TKTTTVNVESKD----NGKKTEVKI---GAKTSVIKEKDGKLVTGK----DKGENGSSTD 164
               ||   : :  |    |  | | ||   ||: :     |   : ||    ||  :| ||
Db        348 -KTKNESAKLSDLEANNAVKGESKITVNGAEYTANAAGDKVTLAGKTMFIDKTASGVSTL 406

Qy        165 EGEGLVTAKE 174
```

```
                |    ||:
Db        407 INEDAAAAKK 416


RESULT 6
F90961
flagellin [imported] - Escherichia coli (strain O157:H7, substrain RIMD 0509952)
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001
C;Accession: F90961
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genc
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: F90961
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-585 <HAY>
A;Cross-references: GB:BA000007; PIDN:BAB36085.1; PID:g13362130; GSPDB:GN00154
A;Experimental source: strain O157:H7, substrain RIMD 0509952
C;Genetics:
A;Gene: ECs2662
C;Superfamily: flagellin

  Query Match             11.1%;  Score 113;  DB 2;  Length 585;
  Best Local Similarity   27.1%;  Pred. No. 1.4;
  Matches   42;  Conservative   27;  Mismatches   62;  Indels   24;  Gaps    7;

Qy         27 LTSVGTEKLSFSANGNKVNITSDTKGLNFAKETAGTNGDTTVHLNGIGSTLTDRAASVKD 86
              |   |: |:   ||  :  || | | ::   :| | ||    ::|| | |:|::||:| |
Db        138 LAKDGSMKIQVGANDGE-TITIDLKKID--SDTLGLNG---FNVNGKG-TITNKAATVSD 190

Qy         87 VLNAGWNIKGVKNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEK 146
              : :||  :     :  ::| :|:        ||    :   || |  | :|        |
Db        191 LTSAGAKLNTTTGLYDLKTENTL-------LTTDAAFDKLGNGDK--VTVGGVDYTYNAK 241

Qy        147 DGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNK 181
              |   | |        :  | |:  | :  |: :|
Db        242 SGDFTTTK--------STAGTGVDAAAQAADSASK 268


RESULT 7
F85809
hypothetical protein fliC [imported] - Escherichia coli (strain O157:H7, substrain EDL93
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C;Accession: F85809
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: F85809
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-585 <STO>
A;Cross-references: GB:AE005174; NID:g12516024; PIDN:AAG56938.1; GSPDB:GN00145; UWGP:Z30
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: fliC
C;Superfamily: flagellin

  Query Match             11.1%;  Score 113;  DB 2;  Length 585;
  Best Local Similarity   27.1%;  Pred. No. 1.4;
  Matches   42;  Conservative   27;  Mismatches   62;  Indels   24;  Gaps    7;

Qy         27 LTSVGTEKLSFSANGNKVNITSDTKGLNFAKETAGTNGDTTVHLNGIGSTLTDRAASVKD 86
              |   |: |:   ||  :  || | | ::   :| | ||    ::|| | |:|::||:| |
Db        138 LAKDGSMKIQVGANDGE-TITIDLKKID--SDTLGLNG---FNVNGKG-TITNKAATVSD 190

Qy         87 VLNAGWNIKGVKNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEK 146
              : :||  :     :  ::| :|:        ||    :   || |  | :|        |
Db        191 LTSAGAKLNTTTGLYDLKTENTL-------LTTDAAFDKLGNGDK--VTVGGVDYTYNAK 241

Qy        147 DGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNK 181
              |   | |        :  | |:  | :  |: :|
Db        242 SGDFTTTK--------STAGTGVDAAAQAADSASK 268


RESULT 8
S61332
IgA-specific metalloendopeptidase (EC 3.4.24.13) homolog - Haemophilus influenzae (st
N;Alternate names: IgA1 protease
C;Species: Haemophilus influenzae
A;Variety: HK635
C;Date: 14-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 08-Dec-2000
C;Accession: S61332
R;Lomholt, H.; Poulsen, K.; Mogens, K.
Mol. Microbiol. 15, 495-506, 1995
A;Title: Comparative characterization of the iga gene encoding IgA1 protease in Neiss
A;Reference number: S61314; MUID:95302961; PMID:7783620
A;Accession: S61332
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-462 <LOM>
A;Cross-references: EMBL:X82488; NID:g732714; PIDN:CAA57871.1; PID:g732715
A;Experimental source: strain HK635
A;Note: the authors did not translate the codon for residue 462
C;Genetics:
A;Gene: iga
C;Superfamily: IgA-specific metalloendopeptidase
C;Keywords: hydrolase; metalloproteinase

  Query Match             11.1%;  Score 112.5;  DB 2;  Length 462;
  Best Local Similarity   25.5%;  Pred. No. 1.2;
  Matches   52;  Conservative   24;  Mismatches   59;  Indels   69;  Gaps   10;

Qy         10 NLKIKQFTYS-LKKDLTDLTSVGTEKLSFSANGN----------KVNITSDTKGLNFAK 57
              |:   |||   | ||   |   |  : ::::|||          :||:  ::|  | :
Db        259 NIYKSQFTKDVLNKDSAGLLK-GNTQYNWTSNGNTSMISNGSELLEVNLFDNSKHTN--R 315

Qy         58 ETAGTNGDTTVHLNGIGSTLTDRAASVKDVLNAG-------------------WNIKGV 97
              | |      |   ||   |||     :|: :| |                   ||  |:
Db        316 EKANYGKSVTFQGNG---TLT-----LKNSINQGAGGLFFEGNYTVEGSSDNIVWNGAGI 367

Qy         98 KNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKDKG 157
                           |: || |  | :  : :    |||  | ::          : ||
Db        368 S--------------VAEGKTVTWKVHNPQSDRL--AKIGKGTLIV----------EGKG 401

Qy        158 ENGSSTDEGEGLVTAKEVIDAVNK 181
              ||  |   |:| |  |:  || ||
Db        402 ENKGSLKVGDGTVILKQQADANNK 425


RESULT 9
A43855
high-molecular-weight surface-exposed protein HMW1 - Haemophilus influenzae
C;Species: Haemophilus influenzae
C;Date: 10-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 08-Oct-1999
C;Accession: A43855
R;Barenkamp, S.J.; Leininger, E.
Infect. Immun. 60, 1302-1313, 1992
A;Title: Cloning, expression, and DNA sequence analysis of genes encoding nontypeable
detella pertussis.
A;Reference number: A43855; MUID:92192797; PMID:1548058
A;Accession: A43855
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1536 <BAR>
A;Cross-references: GB:U08876; GB:M84616; NID:g475770; PIDN:AAA20527.1; PID:g475771
A;Note: sequence inconsistent with the nucleotide translation
A;Note: sequence extracted from NCBI backbone (NCBIN:89235, NCBIP:89239)

  Query Match             11.0%;  Score 112;  DB 2;  Length 1536;
  Best Local Similarity   21.9%;  Pred. No. 5.1;
```

```
  Matches   66;  Conservative   34;  Mismatches   91;  Indels  110;  Gaps   15;

Qy          5 LKAGDNLKIK-----QFTYSLKKDLTDLTSVGTEKLSFSANGN----------KVNITSD 49
              :  | :|: |     :  :|::|||| | : |          |           | ||| :
Db        774 VSTGSSLRFKTSGSTKTGFSIEKDLT-LNATGGNITLLQVEGTDGMIGKGIVAKKNITFE 832

Qy         50 TKGLNFAKETAGT--NGDTTVHLNG----IGSTLTDRAASV---KDV------LNAGWNI 94
                : |    | |   |: |:: |     |||   :    :   |||      | || ||
Db        833 GGNITFGSRKAVTEIEGNVTINNNANVTLIGSDFDNHQKPLTIKKDVIINSGNLTAGGNI 892

Qy         95 ------------------------------KGVKNVDFV------RTYDTVEFLSADTK 117
                                           ||  |:         :  |  : ||  |
Db        893 VNIAGNLTVESNANFKAITNFTFNVGGLFDNKGNSNISIAKGGARFKDIDNSKNLSITTN 952

Qy        118 TTTV-------NVESK-------DNGKKTEVKIGAKTSVIKEKDGKLVTGKDK------- 156
              :::        |: :|       : |  ||::||   |   :|:| |    ||
Db        953 SSSTYRTIISGNITNKNGDLNITNEGSDTEMQIGGDVS---QKEGNLTISSDKINITKQI 1009

Qy        157 -------GENGSSTDEGEGLVTAK----EVIDAVNKAGWRMKTTTAN-------GQTGQA 198
                    |||  |        :| |    ::   :| :|:     ||          | |  |
Db       1010 TIKAGVDGENSDSDATNNANLTIKTKELKLTQDLNISGFNKAEITAKDGSDLTIGNTNSA 1069

Qy        199 D 199
              |
Db       1070 D 1070
```

```
RESULT 10
C48658
flagellin - Escherichia coli (strain U5-41)
C;Species: Escherichia coli
C;Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 26-Aug-1999
C;Accession: C48658
R;Schoenhals, G.; Whitfield, C.
J. Bacteriol. 175, 5395-5402, 1993
A;Title: Comparative analysis of flagellin sequences from Escherichia coli strains posse
A;Reference number: A48658; MUID:93374833; PMID:8366026
A;Accession: C48658
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-584 <SCH>
A;Cross-references: GB:L07388; NID:g290438; PIDN:AAA23798.1; PID:g290439
C;Superfamily: flagellin

  Query Match             10.8%;  Score 109.5;  DB 2;  Length 584;
  Best Local Similarity   27.1%;  Pred. No. 2.5;
  Matches   42;  Conservative   27;  Mismatches   61;  Indels   25;  Gaps    8;

Qy         27 LTSVGTEKLSFSANGNKVNITSDTKGLNFAKETAGTNGDTTVHLNGIGSTLTDRAASVKD 86
              |   |: |:   ||  :  || | | ::   :| | ||    ::|| | |:|::||:| |
Db        138 LAKDGSMKIQVGANDGE-TITIDLKKID--SDTLGLNG---FNVNGKG-TITNKAATVSD 190

Qy         87 VLNAGWNIKGVKNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEK 146
              : :||  :     :  ::| :|:        ||    :   || |  | :|    :  |
Db        191 LTSAGAKLNTTTGLYGLKTENTL-------LTTDAAFDKLGNGDK--VTVGG-VDIYNAK 240

Qy        147 DGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNK 181
               |   | |        :  | |:  | :  |:  |
Db        241 SGDFTTTK--------STAGTGVDAAAQATDSAKK 267
```

```
RESULT 11
I40614
surface array protein - Campylobacter fetus
C;Species: Campylobacter fetus
C;Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 08-Oct-1999
C;Accession: I40614
R;Tummuru, M.K.R.; Blaser, M.J.
Proc. Natl. Acad. Sci. U.S.A. 90, 7265-7269, 1993
A;Title: Rearrangement of sapA homologs with conserved and variable regions in Campyloba
A;Reference number: I40614; MUID:93348254; PMID:8346244
A;Accession: I40614
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-920 <RES>
A;Cross-references: GB:L15800; NID:g289533; PIDN:AAA23019.1; PID:g289534
C;Genetics:
A;Gene: sapA

  Query Match             10.8%;  Score 109.5;  DB 2;  Length 920;
  Best Local Similarity   23.0%;  Pred. No. 4.3;
  Matches   53;  Conservative   33;  Mismatches   81;  Indels   63;  Gaps   11;

Qy          4 TLKAGDNLKIKQFTYSLKKDLTDLTSVGTEKLSF----SANGNKVNIT----SDTKGLNF 55
              || |||          | ||:      |  |:||    ||:|:::::     :||  |
Db        498 TLNAGD--------LDLVKDIDTTGFSGNAKVSFGDTDSADGSQLSVKTGAGNDTLNLEA 549

Qy         56 AKETAGT-----NGDTTVHLNGIGSTLTDRA----------ASVKDVLNAGWNIKGVK-- 98
                |  ||:      |: |: :    | | | |          :| | |:|  ::
Db        550 KKLKAGSLIDGGEGNDTIIMK--ASALADAATLGMIKNIENVTVSDALSANTDVSASSFV 607

Qy         99 NVDFVRTYDTVEFLSADTKTTTVNVESKDNGKK----------------TEVKIGAKTSV 142
              |: :       |     |  |::::||: |                  : : || ::
Db        608 NIGLLADKTDAPFELTVNKNQTIDIQSKEMAKSQILTIKMNDMSGSDDTVNIVLNAK-AI 666

Qy        143 IKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTAN 192
              | : |  :  |       |   |:|          |::||        ||||
Db        667 INQGDKNVAAGAQ--TKGLKIDDG---------IESVNITSVAKDNTTAN 705
```

```
RESULT 12
T39577
hypothetical protein SPBC16E9.02c - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C;Accession: T39577
R;Volckaert, G.; Wood, V.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, August 1997
A;Reference number: Z21865
A;Accession: T39577
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-569 <VOL>
A;Cross-references: EMBL:Z99759; PIDN:CAB16895.1; GSPDB:GN00067; SPDB:SPBC16E9.02c
A;Experimental source: strain 972h-; cosmid c16E9
C;Genetics:
A;Gene: SPDB:SPBC16E9.02c
A;Map position: 2
A;Introns: 33/3; 57/1

  Query Match             10.7%;  Score 108.5;  DB 2;  Length 569;
  Best Local Similarity   29.1%;  Pred. No. 2.9;
  Matches   53;  Conservative   25;  Mismatches   79;  Indels   25;  Gaps    9;

Qy          9 DNLKIKQFTYSLKKD----LTDLTSVGTEKLSFSANGNKVNITSDTKGLNFAKETAGTNG 64
              |  :| : | :    |    | : : || || |   | || ::  | :     |:|  ||
Db        395 DTAEISEKTEAKNADSEANLEEKSDVGEEKESKDEN-NKASLHKDVE----EKDTKITNE 449

Qy         65 DTTVHLNGIGSTLTDRAASVKDVLNAGWNIKGVKNVDFVRTYDTV---EFLSADTKT-TT 120
              ||       | | ||  |   | : |  | |  | |   | | |    |   || |   |
Db        450 DT-------GKTETDVKAKETDSIEA--NDKDEK-TDSKETEDKVEETESKEADVKAKET 499

Qy        121 VNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVN 180
               ::|  |  :||: |  |    ::: | |    |   ::    :|    | :|:| : :
Db        500 DSIEVDDKEEKTDSKETA--DKVEQTDSKDTNEKPAKDDNKEANEKAEKVDSKDVKEKIE 557

Qy        181 KA 182
              :|
Db        558 EA 559
```

```
RESULT 13
H71713
glucose inhibited division protein A (gidA) RP056 - Rickettsia prowazekii
C;Species: Rickettsia prowazekii
C;Date: 21-Nov-1998 #sequence_revision 21-Nov-1998 #text_change 03-Nov-2000
```

```
C;Accession: H71713
R;Andersson, S.G.E.; Zomorodipour, A.; Andersson, J.O.; Sicheritz-Ponten, T.; Alsmark, U
Nature 396, 133-140, 1998
A;Title: The genome sequence of Rickettsia prowazekii and the origin of mitochondria.
A;Reference number: A71630; MUID:99039499; PMID:9823893
A;Accession: H71713
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-621 <AND>
A;Cross-references: GB:AJ235270; GB:AJ235269; NID:g3860572; PIDN:CAA14527.1; PID:g386062
A;Experimental source: strain Madrid E
C;Genetics:
A;Gene: gidA; RP056
C;Superfamily: gidA protein

  Query Match            10.3%;  Score 105;  DB 2;  Length 621;
  Best Local Similarity  27.3%;  Pred. No. 5.6;
  Matches   44;  Conservative   26;  Mismatches   67;  Indels   24;  Gaps    6;

Qy          4 TLKAGDNLKIKQFTYSLKKDLTDLTSVGTEKLSFSANGNKVNITSDTK---GLNFAKETA 60
              |:   :| :: :  |::: |  |                ::|:| :||   || || :
Db        323 TIPGLENAQVLRPGYAIEYDYVD-------------PREINVTLETKKITGLYFAGQIN 368

Qy         61 GTNGDTTVHLNGIGSTLTDRAASVKD----VLNAGWNIKGVKNVDFVRTYDTVEFLSADT 116
              || |       || : : : | ||||    :|    :  ||  :| : |: |||     |
Db        369 GTTGYEEAAGQGIIAGI-NAALSVKDQAPFILTRATSYIGVM-IDDLTTFGTVEPYRMFT 426

Qy        117 KTTTVNVESK-DNGKKTEVKIGAKTSVIKEKDGKLVTGKDK 156
              :   :  : ||      ::| |  || ||  |  | | |
Db        427 SRSEYRLSLRADNADLRLTELGIKIGVITEKRKKFFTKKCK 467


RESULT 14
A56143
surface-array protein homolog sapA2 - Campylobacter fetus (strain 82-40 LP3)
C;Species: Campylobacter fetus
C;Date: 03-Nov-1995 #sequence_revision 03-Nov-1995 #text_change 08-Oct-1999
C;Accession: A56143
R;Dworkin, J.; Tummuru, M.K.R.; Blaser, M.J.
J. Bacteriol. 177, 1734-1741, 1995
A;Title: A lipopolysaccharide-binding domain of the Campylobacter fetus S-layer protein
A;Reference number: A56143; MUID:95204338; PMID:7896695
A;Accession: A56143
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1109 <DWO>
A;Cross-references: GB:S76860; NID:g913763; PIDN:AAB33871.1; PID:g913764
A;Note: this gene appeared to be silent in strain 82-40 LP3
C;Genetics:
A;Gene: sapA2

  Query Match            10.3%;  Score 105;  DB 2;  Length 1109;
  Best Local Similarity  24.0%;  Pred. No. 11;
  Matches   63;  Conservative   27;  Mismatches   90;  Indels   82;  Gaps   13;

Qy          3 NTLKAGDN---------LKIKQFTYSLKKDLTDLTSVG----TEKLSFSANGNKVNITSD 49
              :||  ||          ||:      || :   |   |     |||| : | :  | | |
Db        226 STLNPGDKIDGGAGNDVLKV-----DLKNNFKGLKDDGYIKNIEKLSLT-NSSVSNRTFD 279

Qy         50 TKGLNFAKETAGTNGD-----------TTVHLNGIGST----------LTDRAASVKDV 87
               ||::   :|   :|:             | :||   |           : | :|   ||
Db        280 AKGID-GLQTVALSGEKGISVTNLANIVDVEVNGFKGTNFNVDSIYADKVLDGSA---DV 335

Qy         88 LNAGWNIKGVKNVDFVRTYDTVE----------FLSAD----------------TKTT 119
              |   |  | |      | | :|          |:|||                  |||
Db        336 QNLKVNGVGAKGASVAITADKIETLNLNTTGSQSFVSADVASISVKGNANLSLATGAKTT 395

Qy        120 TVNVESKDNGKKTEVKIGAKTSVIKEKDGK-LVTGKDKGEN-----GSSTDEGEGLVTAK 173
              |::  |       ::   |  : ||  :|   :| ||   |     |:  ||   ||
Db        396 TLDASSFGGALDADLSTSASVTSIKGGNGNDKITIKDVAVNVAIDGGAGNDE---LVIKG 452

Qy        174 EVIDAVNKAGWRMKTTTANGQT 195
                 | :      ::  | :| |
Db        453 STADTLQPTLTNIEKVTIDGNT 474


RESULT 15
B43855
high-molecular-weight surface-exposed protein - Haemophilus influenzae
C;Species: Haemophilus influenzae
C;Date: 10-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 18-Nov-1994
C;Accession: B43855
R;Barenkamp, S.J.; Leininger, E.
Infect. Immun. 60, 1302-1313, 1992
A;Title: Cloning, expression, and DNA sequence analysis of genes encoding nontypeable
detella pertussis.
A;Reference number: A43855; MUID:92192797; PMID:1548058
A;Accession: B43855
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1477 <BAR>
A;Note: sequence inconsistent with the nucleotide translation
A;Note: sequence extracted from NCBI backbone (NCBIN:89237, NCBIP:89240)

  Query Match            10.3%;  Score 105;  DB 2;  Length 1477;
  Best Local Similarity  28.3%;  Pred. No. 15;
  Matches   58;  Conservative   23;  Mismatches   76;  Indels   48;  Gaps   12;

Qy          4 TLKAGDNLKIKQFTYSLKKDLTDLTSVGTEKLSFSANGNKVNITSDTKGLNFAK--ETAG 61
              | |:|  :: |    | : ::|  |  ||  :  : :|| ||:|::   |      |
Db       1191 TTKSGSKIEAK----SGEANVTSAT--GT--IGGTISGNTVNVTANAGDLTVGNGAEINA 1242

Qy         62 TNGDTTVHLNGIGSTLTDRAAS-------VKDVLNAGWNIKGVKNVDFVRTYDTVEFLSA 114
              | |  |  |   |:||| | |          |:|    :| |  |   |         :
Db       1243 TEGAAT--LTATGNTLTTEAGSSITSTKGQVDLLAQNGSIAGSINAANV---------TL 1291

Qy        115 DTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKDKGENGSSTD------EGEG 168
              :|  |   |   |      :|  : | ||  || ||    :  :| ||:       | |
Db       1292 NTTGTLTTVAGSD------IKATSGTLVINAKDAKL----NGDASGDSTEVNAVNASGSG 1341

Qy        169 LVTAKEVIDAVNKAGWRMKTTTANG 193
               |||     :||  |      | ||
Db       1342 SVTA-ATSSSVNITG---DLNTVNG 1362


Search completed: October  6, 2003, 09:33:42
Job time : 7.75367 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: October 6, 2003, 09:06:20 ; Search time 3.10395 Seconds
(without alignments)
3045.266 Million cell updates/sec

Title: US-09-771-382-39
Perfect score: 1018
Sequence: 1 SANTLKAGDNLKIKQFTYSL.....AGWRMKTTTANGQTGQADKF 201

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	116.5	11.4	507	1	FLIC_SALON	Q06974 salmonella
2	105	10.3	621	1	GIDA_RICPR	Q9ze90 rickettsia
3	103	10.1	504	1	FLIC_SALSE	Q06983 salmonella
4	102	10.0	504	1	FLIC_SALBU	Q06969 salmonella
5	100.5	9.9	550	1	FLIC_SHIFL	Q08860 shigella fl
6	99	9.7	918	1	YMJB_CAEEL	P34487 caenorhabdi
7	99	9.7	1005	1	Y456_CHLTR	O84462 chlamydia t
8	98.5	9.7	798	1	YC00_MYCPN	Q50288 mycoplasma
9	98.5	9.7	936	1	PMP7_CHLPN	Q9z898 chlamydia p
10	98	9.6	444	1	SLAP_LACAC	P35829 lactobacill
11	97	9.5	504	1	FLIC_SALMC	Q06981 salmonella
12	96.5	9.5	505	1	FLJB_SALTY	P52616 salmonella
13	96.5	9.5	2334	1	WAPA_BACSU	Q07833 bacillus su
14	96	9.4	504	1	FLIC_SALMO	Q06973 salmonella
15	95	9.3	1390	1	RPOB_MYCGE	P47583 mycoplasma
16	94.5	9.3	1608	1	HLYA_SERMA	P15320 serratia ma
17	94	9.2	874	1	SLAP_BACLI	P49052 bacillus li
18	94	9.2	1032	1	KINN_HUMAN	Q12840 homo sapien
19	94	9.2	1592	1	GTF2_STRDO	P27470 streptococc
20	93.5	9.2	474	1	FLDB_PSEAE	Q9k3c5 pseudomonas
21	93.5	9.2	1007	1	Y741_CHLMU	Q9pjt6 chlamydia m
22	93.5	9.2	1694	1	IGA0_HAEIN	P44969 haemophilus
23	93.5	9.2	1702	1	IGA2_HAEIN	P45384 haemophilus
24	93	9.1	365	1	FLC1_PROMI	P42272 proteus mir
25	93	9.1	787	1	YC88_MYCPN	P75489 mycoplasma
26	93	9.1	1186	1	CAGA_HELPY	P55980 helicobacte
27	93	9.1	1435	1	EBA1_PLAFC	P19214 plasmodium
28	92.5	9.1	370	1	PGL1_PENOL	Q9y834 penicillium
29	92.5	9.1	494	1	FLIC_SALTY	P06179 salmonella
30	92.5	9.1	501	1	FLAA_AQUPY	P46210 aquifex pyr
31	92.5	9.1	1300	1	120K_RICRI	P14914 rickettsia
32	92.5	9.1	1654	1	OMPB_RICRI	Q53047 r outer mem
33	92	9.0	504	1	FLIC_SALEN	Q06972 salmonella
34	92	9.0	504	1	FLIC_SALNA	O52959 salmonella
35	92	9.0	935	1	EAE_ECO11	O31000 escherichia
36	92	9.0	1167	1	CAGA_HELPJ	Q9zlt1 helicobacte
37	92	9.0	1325	1	YDEK_ECOLI	P32051 escherichia
38	91.5	9.0	497	1	EXU1_DROPS	Q24618 drosophila
39	91.5	9.0	1541	1	IGA1_HAEIN	P42782 haemophilus
40	91.5	9.0	1645	1	OMPB_RICTY	P96989 r outer mem
41	91	8.9	307	1	K1PF_BORBU	O51575 borrelia bu
42	91	8.9	414	1	YOJL_BACSU	O31852 bacillus su
43	91	8.9	504	1	FLIC_SALDU	Q06971 salmonella
44	91	8.9	536	1	NAB2_RALSO	Q8xqg4 ralstonia s
45	91	8.9	1569	1	YPJA_ECOLI	P52143 escherichia

ALIGNMENTS

RESULT 1
FLIC_SALON
ID FLIC_SALON STANDARD; PRT; 507 AA.
AC Q06974;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Flagellin (Phase-1-C flagellin).
GN FLIC.
OS Salmonella oranienberg.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=28147;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 9239;
RX MEDLINE=93374829; PubMed=7690024;
RA Masten B.J., Joys T.M.;
RT "Molecular analyses of the Salmonella g. . . flagellar antigen
RT complex.";
RL J. Bacteriol. 175:5359-5365(1993).
CC -!- FUNCTION: FLAGELLIN IS THE SUBUNIT PROTEIN WHICH POLYMERIZES TO
CC FORM THE FILAMENTS OF BACTERIAL FLAGELLA.
CC -!- MISCELLANEOUS: INDIVIDUAL SALMONELLA SEROTYPES USUALLY ALTERNATE
CC BETWEEN THE PRODUCTION OF 2 ANTIGENIC FORMS OF FLAGELLA, TERMED
CC PHASE-1 AND PHASE-2, EACH SPECIFIED BY SEPARATE STRUCTURAL GENES.
CC -!- SIMILARITY: BELONGS TO THE BACTERIAL FLAGELLIN FAMILY.
CC ----------

CC ----------
DR EMBL; Z15070; CAA78779.1; -.
DR PIR; A53465; A53465.
DR PIR; S33192; S33192.
DR InterPro; IPR001029; Flagellin_C.
DR InterPro; IPR001492; FlagellinN.
DR Pfam; PF00700; Flagellin_C; 1.
DR Pfam; PF00669; Flagellin_N; 1.
DR PRINTS; PR00207; FLAGELLIN.
KW Flagella.
FT INIT_MET 0 0 BY SIMILARITY.
SQ SEQUENCE 507 AA; 53065 MW; 17A88B4BBA4A6125 CRC64;

Query Match 11.4%; Score 116.5; DB 1; Length 507;
Best Local Similarity 31.1%; Pred. No. 0.39;
Matches 59; Conservative 27; Mismatches 73; Indels 31; Gaps 11;

```
Qy          4 TLKAGDNLKIKQFTYSLKKDLTDLTSVGTEKLSFSANGNKVNITSDTKGLNFAKET-AGT 62
              |  | :|  :  |  :     ||      : :: |  | ||  | | ||::|  :| ||
Db        238 TADAQNNTAVDLFKSTKSAAGTD----DAKAIATSIKGGKVGDTFDYKGVSFTIDTKAGD 293
```

```
Qy          63 NGDTTVH--LNG--IGSTLTDRAASVKDVLNAGWNIKGVKNVDFVRTYDTV---EFLSAD 115
               :|: ||   :||  :  |::|  ||  || :|   |:  |:|     | :|   :|  ||
Db         294 DGNGTVSTTINGEKVTLTISDIGASATDVNSA--KIQSSKDV-----YTSVVSGQFTFAD 346

Qy         116 TKTTTVNVESKD----NGKKTEVKI---GAKTSVIKEKDGKLVTGK----DKGENGSSTD 164
                ||   : :  |    |  | | ||   ||: :     |   : ||    ||  :| ||
Db         347 -KTKNESAKLSDLEANNAVKGESKITVNGAEYTANAAGDKVTLAGKTMFIDKTASGVSTL 405

Qy         165 EGEGLVTAKE 174
                 |    ||:
Db         406 INEDAAAAKK 415


RESULT 2
GIDA_RICPR
ID   GIDA_RICPR     STANDARD;      PRT;   621 AA.
AC   Q9ZE90;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Glucose inhibited division protein A.
GN   GIDA OR RP056.
OS   Rickettsia prowazekii.
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC   Rickettsiaceae; Rickettsieae; Rickettsia.
OX   NCBI_TaxID=782;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Madrid E;
RX   MEDLINE=99039499; PubMed=9823893;
RA   Andersson S.G.E., Zomorodipour A., Andersson J.O.,
RA   Sicheritz-Ponten T., Alsmark U.C.M., Podowski R.M., Naeslund A.K.,
RA   Eriksson A.-S., Winkler H.H., Kurland C.G.;
RT   "The genome sequence of Rickettsia prowazekii and the origin of
RT   mitochondria.";
RL   Nature 396:133-140(1998).
CC   -!- FUNCTION: NOT KNOWN.
CC   -!- SIMILARITY: Belongs to the gidA family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AJ235270; CAA14527.1; -.
DR   PIR; H71713; H71713.
DR   HAMAP; MF_00129; -; 1.
DR   InterPro; IPR001327; FAD_pyr_redox.
DR   InterPro; IPR002218; GIDA.
DR   InterPro; IPR004416; GidA_sub.
DR   InterPro; IPR001100; Pyr_redox.
DR   Pfam; PF01134; GIDA; 1.
DR   PRINTS; PR00368; FADPNR.
DR   PRINTS; PR00411; PNDRDTASEI.
DR   ProDom; PD003738; GIDA; 1.
DR   TIGRFAMs; TIGR00136; gidA; 1.
DR   PROSITE; PS01280; GIDA_1; 1.
DR   PROSITE; PS01281; GIDA_2; 1.
KW   Complete proteome.
SQ   SEQUENCE   621 AA;  69195 MW;  7379A85B4E6ACADF CRC64;

  Query Match             10.3%;  Score 105;  DB 1;  Length 621;
  Best Local Similarity   27.3%;  Pred. No. 3;
  Matches   44;  Conservative   26;  Mismatches   67;  Indels   24;  Gaps    6;

Qy           4 TLKAGDNLKIKQFTYSLKKDLTDLTSVGTEKLSFSANGNKVNITSDTK---GLNFAKETA 60
               |:   :| :: :  |::: |  |               ::|:| :||   || || :
Db         323 TIPGLENAQVLRPGYAIEYDYVD-------------PREINVTLETKKITGLYFAGQIN 368

Qy          61 GTNGDTTVHLNGIGSTLTDRAASVKD----VLNAGWNIKGVKNVDFVRTYDTVEFLSADT 116
                  || |       || : : : | ||||    :|    :  ||  :| : |: |||     |
Db         369 GTTGYEEAAGQGIIAGI-NAALSVKDQAPFILTRATSYIGVM-IDDLTTFGTVEPYRMFT 426

Qy         117 KTTTVNVESK-DNGKKTEVKIGAKTSVIKEKDGKLVTGKDK 156
                 :   : : ||      ::| |  || ||  |  | | |
Db         427 SRSEYRLSLRADNADLRLTELGIKIGVITEKRKKFFTKKCK 467


RESULT 3
FLIC_SALSE
ID   FLIC_SALSE     STANDARD;      PRT;   504 AA.
AC   Q06983;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   Flagellin (Phase-1-C flagellin).
GN   FLIC.
OS   Salmonella senftenberg.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Salmonella.
OX   NCBI_TaxID=28150;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93374829; PubMed=7690024;
RA   Masten B.J., Joys T.M.;
RT   "Molecular analyses of the Salmonella g. . . flagellar antigen
RT   complex.";
RL   J. Bacteriol. 175:5359-5365(1993).
CC   -!- FUNCTION: FLAGELLIN IS THE SUBUNIT PROTEIN WHICH POLYMERIZES TO
CC       FORM THE FILAMENTS OF BACTERIAL FLAGELLA.
CC   -!- MISCELLANEOUS: INDIVIDUAL SALMONELLA SEROTYPES USUALLY ALTERNATE
CC       BETWEEN THE PRODUCTION OF 2 ANTIGENIC FORMS OF FLAGELLA, TERMED
CC       PHASE-1 AND PHASE-2, EACH SPECIFIED BY SEPARATE STRUCTURAL GENES.
CC   -!- SIMILARITY: BELONGS TO THE BACTERIAL FLAGELLIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z15072; CAA78781.1; -.
DR   PIR; S33194; S33194.
DR   InterPro; IPR001029; Flagellin_C.
DR   InterPro; IPR001492; FlagellinN.
DR   Pfam; PF00700; Flagellin_C; 1.
DR   Pfam; PF00669; Flagellin_N; 1.
DR   PRINTS; PR00207; FLAGELLIN.
KW   Flagella.
FT   INIT_MET      0      0       BY SIMILARITY.
SQ   SEQUENCE   504 AA;  52864 MW;  BFD14C4F125C2BAC CRC64;

  Query Match             10.1%;  Score 103;  DB 1;  Length 504;
  Best Local Similarity   30.8%;  Pred. No. 3.3;
  Matches   53;  Conservative   27;  Mismatches   62;  Indels   30;  Gaps   11;

Qy          25 TDLTSVGTEK---LSFSANGNKVNITSDTKGLNFAKET-AGTNGDTTVH--LNGIGSTLT 78
               |  :: ||::   :: :  | |   | | ||::|  :| || :|: ||   :||   |||
Db         249 TTKSTAGTDEAKAIASAIKGGKEGDTFDYKGVSFTIDTKAGNDGNGTVSTTINGEKVTLT 308

Qy          79 --DRAASVKDVLNAGWNIKGVKNVDFVRTYDTV---EFLSADTKTTTVNVESKD----NG 129
                 |  |   :| :|   ::  |||      | :|   :| : | ||   : :  |    |
Db         309 VADITAGAANVNDA--TLQSSKNV-----YTSVVNGQF-TFDDKTKNESAKLSDLEANNA 360

Qy         130 KKTEVKI---GAKTSVIKEKDGKLVTGK----DKGENGSSTDEGEGLVTAKE 174
                | | ||   ||: :     |   : ||    ||  :| ||   |    ||:
Db         361 VKGESKITVNGAEYTANAAGDKVTLAGKTMFIDKTASGVSTLINEDAAAAKK 412


RESULT 4
FLIC_SALBU
```

```
ID   FLIC_SALBU     STANDARD;      PRT;   504 AA.
AC   Q06969;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   Flagellin (Phase-1-C flagellin).
GN   FLIC.
OS   Salmonella budapest.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Salmonella.
OX   NCBI_TaxID=28143;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93374829; PubMed=7690024;
RA   Masten B.J., Joys T.M.;
RT   "Molecular analyses of the Salmonella g. . . flagellar antigen
RT   complex.";
RL   J. Bacteriol. 175:5359-5365(1993).
CC   -!- FUNCTION: FLAGELLIN IS THE SUBUNIT PROTEIN WHICH POLYMERIZES TO
CC       FORM THE FILAMENTS OF BACTERIAL FLAGELLA.
CC   -!- MISCELLANEOUS: INDIVIDUAL SALMONELLA SEROTYPES USUALLY ALTERNATE
CC       BETWEEN THE PRODUCTION OF 2 ANTIGENIC FORMS OF FLAGELLA, TERMED
CC       PHASE-1 AND PHASE-2, EACH SPECIFIED BY SEPARATE STRUCTURAL GENES.
CC   -!- SIMILARITY: BELONGS TO THE BACTERIAL FLAGELLIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z15065; CAA78774.1; -.
DR   PIR; S33186; S33186.
DR   InterPro; IPR001029; Flagellin_C.
DR   InterPro; IPR001492; FlagellinN.
DR   Pfam; PF00700; Flagellin_C; 1.
DR   Pfam; PF00669; Flagellin_N; 1.
DR   PRINTS; PR00207; FLAGELLIN.
KW   Flagella.
FT   INIT_MET      0      0       BY SIMILARITY.
SQ   SEQUENCE   504 AA;  52790 MW;  1FD1498751B6475E CRC64;

  Query Match             10.0%;  Score 102;  DB 1;  Length 504;
  Best Local Similarity   30.8%;  Pred. No. 3.8;
  Matches   53;  Conservative   27;  Mismatches   62;  Indels   30;  Gaps   11;

Qy         25 TDLTSVGTEK---LSFSANGNKVNITSDTKGLNFAKET-AGTNGDTTVH--LNGIGSTLT 78
              |  :: ||::   :: :  | |   | | ||::|  :| || :|: ||   :||   |||
Db        249 TTKSTAGTDEAKAIAGAIKGGKEGDTFDYKGVSFTIDTKAGNDGNGTVSTTINGEKVTLT 308

Qy         79 --DRAASVKDVLNAGWNIKGVKNVDFVRTYDTV---EFLSADTKTTTVNVESKD----NG 129
                |  |   :| :|   ::  |||     | :|   :| : | ||   : :  |    |
Db        309 VADITAGAANVNDA--TLQSSKNV-----YTSVVNGQF-TFDDKTKNESAKLSDLEANNA 360

Qy        130 KKTEVKI---GAKTSVIKEKDGKLVTGK----DKGENGSSTDEGEGLVTAKE 174
               | | ||   ||: :      |   : ||    ||  :| ||   |    ||:
Db        361 VKGESKITVNGAEYTANAAGDKVTLAGKTMFIDKTASGVSTLINEDAAAAKK 412


RESULT 5
FLIC_SHIFL
ID   FLIC_SHIFL     STANDARD;      PRT;   550 AA.
AC   Q08860;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Flagellin.
GN   FLIC OR SF1966.
OS   Shigella flexneri.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Shigella.
OX   NCBI_TaxID=623;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=IID642;
RX   MEDLINE=94335647; PubMed=8057852;
RA   Tominaga A., Mahmoud M.A.-H., Mukaihara T., Enomoto M.;
RT   "Molecular characterization of intact, but cryptic, flagellin genes
RT   in the genus Shigella.";
RL   Mol. Microbiol. 12:277-285(1994).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=301 / Serotype 2a;
RX   MEDLINE=22272406; PubMed=12384590;
RA   Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
RA   Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
RA   Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
RA   Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA   Yu J.;
RT   "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
RT   through comparison with genomes of Escherichia coli K12 and O157.";
RL   Nucleic Acids Res. 30:4432-4441(2002).
CC   -!- FUNCTION: FLAGELLIN IS THE SUBUNIT PROTEIN WHICH POLYMERIZES TO
CC       FORM THE FILAMENTS OF BACTERIAL FLAGELLA.
CC   -!- SIMILARITY: BELONGS TO THE BACTERIAL FLAGELLIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D16819; BAA04093.1; -.
DR   EMBL; AE015215; AAN43516.1; -.
DR   PIR; S44980; S44980.
DR   InterPro; IPR001029; Flagellin_C.
DR   InterPro; IPR001492; FlagellinN.
DR   Pfam; PF00700; Flagellin_C; 1.
DR   Pfam; PF00669; Flagellin_N; 1.
DR   PRINTS; PR00207; FLAGELLIN.
KW   Flagella.
SQ   SEQUENCE   550 AA;  56636 MW;  CC921C9A8EF200B6 CRC64;

  Query Match              9.9%;  Score 100.5;  DB 1;  Length 550;
  Best Local Similarity   25.8%;  Pred. No. 5.4;
  Matches   57;  Conservative   29;  Mismatches   96;  Indels   39;  Gaps   11;

Qy          4 TLKAGDNLKIKQFTYSLKKDLTDLTSVGTEKLSFSANGNKVNITSDTKGLNFAKETAGTN 63
              |   || | |   |  | |: :  :       : :||  |   | |  | : :     |
Db        299 TASNGDKLYI-DTTGRLTKNGSGASLTEASLSTLAANNTKAT-TIDIGGTSISFTGNSTT 356

Qy         64 GDT-TVHLNG--IGSTLTDRAASVK----DVLNAGWNIKG-----VKNVDFVRTYDTVEF 111
               || |  : |  :      |:| |       |   ||::: |      | || |     |:
Db        357 PDTITYSVTGAKVDQAAFDKAVSTSGNNVDFTTAGYSVNGTTGAVTKGVDSV----YVDN 412

Qy        112 LSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVT 171
              | | : ||:   :|:|  |    |:  :| |: |||| |  |     ::|   : |
Db        413 NEALTTSDTVDFYLQDDGSVTN---GSGKAVYKDADGKLTT--DAETKAATT--ADPLKA 465

Qy        172 AKEVIDAVNK-------------AGWRMKTTTANGQTGQA 198
                | | :::|             |   :  || |    |:
Db        466 LDEAISSIDKFRSSLGAVQNRLDSAVTNLNNTTTNLSEAQS 506


RESULT 6
YMJB_CAEEL
ID   YMJB_CAEEL     STANDARD;      PRT;   918 AA.
AC   P34487;
DT   01-FEB-1994 (Rel. 28, Created)
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Hypothetical protein F59B2.12 in chromosome III.
```

```
GN   F59B2.12.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RX   MEDLINE=94150718; PubMed=7906398;
RA   Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA   Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA   Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,
RA   Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
RA   Johnston L., Jones M., Kershaw J., Kirsten J., Laisster N.,
RA   Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
RA   Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Shownkeen R.,
RA   Sims M., Smaldon N., Smith A., Smith M., Sonnhammer E., Staden R.,
RA   Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
RA   Waterston R., Watson A., Weinstock L., Wilkinson-Sproat J.,
RA   Wohldman P.;
RT   "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT   elegans.";
RL   Nature 368:32-38(1994).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z11505; CAA77581.1; -.
DR   PIR; G88545; G88545.
DR   WormPep; F59B2.12; CE01024.
KW   Hypothetical protein.
SQ   SEQUENCE   918 AA;  96560 MW;  E464FD86B14945DE CRC64;

  Query Match             9.7%;  Score 99;  DB 1;  Length 918;
  Best Local Similarity   21.3%;  Pred. No. 12;
  Matches   54;  Conservative   32;  Mismatches   84;  Indels   84;  Gaps    9;

Qy          6 KAGDNLKIKQFTYSL-------KKDLTDLTSVGTEKLS----------FSANGNKVNITS 48
              :| :| |||    |:        |   | :| | || |          |:| ||:|  |
Db        118 EANENTKIKSADGSVIETGKSHNKSSDDASSYGLEKSSKTYADKNGTMLSSNTNKINNQS 177

Qy         49 DTKGLN--------------FAKETAG--------------TNGDTTVHLNGIGSTLTDR 80
                :  |:              | :   |              |   ::  :|    :|:|
Db        178 RSAALDEGNEFVNQQNADGTFLRNNTGHKNTDEHLSHNVLDENAQMSIGADGTSHNITNR 237

Qy         81 AASVKDVLNAGWNIKGVKNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKT 140
               || |  ||  :            :   | | |       | :|::  ||     |:
Db        238 KGSVGDSHNAASD----------AHSNFESLDAQG-----NKKSQNYSKKAASASGSNA 281

Qy        141 SV------IKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEV-------------IDAVNK 181
                    :|  ||  ::      | :| |:     ||:||             ::| :
Db        282 DFESNLESLKNADGTSMSNSTGNFNNTSYDK----ATAEEVMSKKNVNADGTSSMEASHA 337

Qy        182 AGWRMKTTTANGQT 195
                  |  :|:||:
Db        338 GSNSSKINSASGQS 351


RESULT 7
Y456_CHLTR
ID   Y456_CHLTR     STANDARD;      PRT;  1005 AA.
AC   O84462;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Protein CT456 precursor.
GN   CT456.
OS   Chlamydia trachomatis.
OC   Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX   NCBI_TaxID=813;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=D/UW-3/Cx;
RX   MEDLINE=99000809; PubMed=9784136;
RA   Stephens R.S., Kalman S., Lammel C.J., Fan J., Marathe R., Aravind L.,
RA   Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V.,
RA   Davis R.W.;
RT   "Genome sequence of an obligate intracellular pathogen of humans:
RT   Chlamydia trachomatis.";
RL   Science 282:754-759(1998).
CC   -!- SIMILARITY: BELONGS TO THE CHLAMYDIAL CPN0572/CT456/TC0741
CC       FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE001319; AAC68056.1; -.
DR   PIR; C71513; C71513.
DR   PHCI-2DPAGE; O84462; -.
KW   Signal; Complete proteome.
FT   SIGNAL       1     40       POTENTIAL.
FT   CHAIN       41   1005       PROTEIN CT456.
SQ   SEQUENCE   1005 AA;  102131 MW;  EC47EC389851CD1E CRC64;

  Query Match             9.7%;  Score 99;  DB 1;  Length 1005;
  Best Local Similarity   22.6%;  Pred. No. 13;
  Matches   48;  Conservative   25;  Mismatches   71;  Indels   68;  Gaps   11;

Qy         41 GNKVNITSDTKGLNFAKETAGTNGDTTVHLNGIGSTLTDRAASVKDVLNAGWNIKGVK-- 98
              | :   :|:  |::|: || |          |  :  :   :|:  ::|   |: |
Db        490 GTQAGPSSEDDGISFSNETPGA---------GPAAAPSPTPSSI-PIINVNVNVGGTNVN 539

Qy         99 ----NVDFVRTYDTVEFLSADT------------KTTTVNVESKD-------NG--KKTE 133
                  ||:   |  | :   | |             :| :|   ||       ||   :||
Db        540 IGDTNVNTTNTTPTTQSTDASTDTSDIDDINTNNQTDDINTTDKDSDGAGGVNGDISETE 599

Qy        134 VKIGAKTSVI--KEKDGKLVTGKD------------------KGENGSSTDEGEGLVTA 172
                  |  :  :   | |     | |                    ||| ||   || : |
Db        600 SSSGDDSGSVSSSESDKNASVGNDGPAMKDILSAVRKHLDVVYPGENGGST---EGPLPA 656

Qy        173 KEVI-----DAVNKAGWRMKTTTANGQTGQAD 199
                : :     |  ||     : |  :| ||  |
Db        657 NQTLGDVISDVENKGS--AQDTKLSGNTGAGD 686


RESULT 8
YC00_MYCPN
ID   YC00_MYCPN     STANDARD;      PRT;   798 AA.
AC   Q50288;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Hypothetical lipoprotein MPN200 precursor (GT9_orf798).
GN   MPN200 OR MP631.
OS   Mycoplasma pneumoniae.
OC   Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX   NCBI_TaxID=2104;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 29342 / M129;
RX   MEDLINE=96177562; PubMed=8604303;
RA   Hilbert H., Himmelreich R., Plagens H., Herrmann R.;
RT   "Sequence analysis of 56 kb from the genome of the bacterium
RT   Mycoplasma pneumoniae comprising the dnaA region, the atp operon and a
RT   cluster of ribosomal protein genes.";
RL   Nucleic Acids Res. 24:628-639(1996).
```

```
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 29342 / M129;
RX   MEDLINE=97105885; PubMed=8948633;
RA   Himmelreich R., Hilbert H., Plagens H., Pirkl E., Li B.-C.,
RA   Herrmann R.;
RT   "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT   pneumoniae.";
RL   Nucleic Acids Res. 24:4420-4449(1996).
CC   -!- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor
CC       (Potential).
CC   -!- SIMILARITY: BELONGS TO THE MG185 / MG260 FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U34795; AAC43681.1; -.
DR   EMBL; AE000060; AAB96279.1; -.
DR   PIR; S62791; S62791.
DR   InterPro; IPR004890; Lipoprotein_10.
DR   InterPro; IPR004984; Lipoprotein_X.
DR   Pfam; PF03202; Lipoprotein_10; 1.
DR   Pfam; PF03305; Lipoprotein_X; 1.
DR   PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
KW   Hypothetical protein; Lipoprotein; Membrane; Signal;
KW   Complete proteome.
FT   SIGNAL        1     22       POTENTIAL.
FT   CHAIN        23    798       HYPOTHETICAL LIPOPROTEIN MPN200.
FT   LIPID        23     23       N-ACYL DIGLYCERIDE (POTENTIAL).
SQ   SEQUENCE   798 AA;  87628 MW;  3E2471D7EF0A3CA0 CRC64;

  Query Match             9.7%;  Score 98.5;  DB 1;  Length 798;
  Best Local Similarity   21.1%;  Pred. No. 11;
  Matches   51;  Conservative   33;  Mismatches   95;  Indels   63;  Gaps    9;

Qy          7 AGDNLKIKQFTYSLKKDLTDLTSVGTEKLSFSANGNKVNITSDTKGLNFAKETAGTNGDT 66
              |  |     | | ::    | ::   :  |:    |  |: :| | |         ||
Db        337 AAGNSDYNNFFYKVQNGRADFSNFNNKGSSYK---NLKNVFNDYKNL------IAQNG-- 385

Qy         67 TVHLNGIGSTLTD------RAASVKDVLNAGWNIKGVKNVDFVRTYDTVEFLSADTKTTT 120
               :::|  ||  ::      | |:      ::  |  :  |  | |   |:   : ||
Db        386 -LYVNKGGSYSSNFQKFHQLAYSISSTSGFAYSFAGQNSKRFKFT-DDGTFVEYPSYTTE 443

Qy        121 VNVESKDNG----------------------KKTEVKIGAKTSVIKEK--DGKLVTGKD 155
              ||    :||                      :::|:  ||    :|  | |
Db        444 VNAPESNNGNDGKQQGQSDQGNLLGTFEVVDKSTSDIKVKPKTQAESKKSSDSKQTANTG 503

Qy        156 KGENGSS--------------TDEGEG----LVTAKEVIDAVNKAGWRMKTTTANGQTG 196
              || |                  |: |     |||  |:|  : ||  : : | |:|::
Db        504 KGSNSKQQTPKKTISLYKTKIPQDKTENVDAFLVTDSELISKLEKAKNKKEETKASGKSA 563

Qy        197 QA 198
                |
Db        564 SA 565


RESULT 9
PMP7_CHLPN
ID   PMP7_CHLPN     STANDARD;      PRT;   936 AA.
AC   Q9Z898; Q9JS42; Q9Z5Q4;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Probable outer membrane protein pmp7 precursor (Polymorphic membrane
DE   protein 7) (Outer membrane protein 12).
GN   PMP7 OR OMP12 OR CPN0445 OR CP0308.
OS   Chlamydia pneumoniae (Chlamydophila pneumoniae).
OC   Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX   NCBI_TaxID=83558;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CWL029;
RX   MEDLINE=99206606; PubMed=10192388;
RA   Kalman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,
RA   Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
RT   "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";
RL   Nat. Genet. 21:385-389(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AR39;
RX   MEDLINE=20150255; PubMed=10684935;
RA   Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA   White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,
RA   Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA   Gwinn M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
RA   Eisen J., Fraser C.M.;
RT   "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
RT   pneumoniae AR39.";
RL   Nucleic Acids Res. 28:1397-1406(2000).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=J138;
RX   MEDLINE=20330349; PubMed=10871362;
RA   Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
RA   Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
RT   "Comparison of whole genome sequences of Chlamydia pneumoniae J138
RT   from Japan and CWL029 from USA.";
RL   Nucleic Acids Res. 28:2311-2314(2000).
RN   [4]
RP   SEQUENCE OF 658-936 FROM N.A.
RC   STRAIN=VR1310;
RX   MEDLINE=20007584; PubMed=10539856;
RA   Christiansen G., Boesen T., Hjerno K., Daugaard L., Mygind P.,
RA   Madsen A.S., Knudsen K., Falk E., Birkelund S.;
RT   "Molecular biology of Chlamydia pneumoniae surface proteins and their
RT   role in immunopathogenicity.";
RL   Am. Heart J. 138:S491-S495(1999).
CC   -!- SUBCELLULAR LOCATION: CELL WALL SURFACE (ELEMENTARY BODIES)
CC       (POTENTIAL).
CC   -!- SIMILARITY: BELONGS TO THE PMP OUTER MEMBRANE PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE001627; AAD18589.1; -.
DR   EMBL; AE002193; AAF38165.1; -.
DR   EMBL; AP002546; BAA98653.1; -.
DR   EMBL; AJ133034; CAB37067.1; -.
DR   PIR; B81591; B81591.
DR   PIR; C72078; C72078.
DR   PIR; C86546; C86546.
DR   PHCI-2DPAGE; Q9Z898; -.
DR   TIGR; CP0308; -.
DR   InterPro; IPR006315; Autotransport.
DR   InterPro; IPR005546; Autotransporter.
DR   InterPro; IPR003368; Chlamydia_PMP.
DR   Pfam; PF03797; Autotransporter; 1.
DR   Pfam; PF02415; DUF145; 2.
DR   TIGRFAMs; TIGR01414; autotrans_barl; 1.
DR   TIGRFAMs; TIGR01376; POMP_repeat; 6.
KW   Outer membrane; Signal; Multigene family; Complete proteome.
FT   SIGNAL        1     23       POTENTIAL.
FT   CHAIN        24    936       PROBABLE OUTER MEMBRANE PROTEIN PMP7.
FT   CONFLICT    658    666       PTRHGFRHI -> EDNIRYRHN (IN REF. 4).
FT   CONFLICT    822    822       Y -> H (IN REF. 1 AND 4).
SQ   SEQUENCE   936 AA;  100105 MW;  3981DB3C950AF95A CRC64;
```

```
  Query Match             9.7%;  Score 98.5;  DB 1;  Length 936;
  Best Local Similarity   24.3%;  Pred. No. 13;
  Matches   59;  Conservative   32;  Mismatches   71;  Indels   81;  Gaps   13;

Qy         27 LTSVGTEKLSFSAN----------GNKVNITSDTKGL---NFAKETAG------------ 61
              :|  | : |||: |          | ||:|::    |   | :  :||
Db        274 VTLTGNKNLSFTNNTALTYGGAISGLKVSISAGGPTLFQSNISGSSAGQGGGGAINIASA 333

Qy         62 -------TNGDTTVHLNGI--GSTLTDRAASVKDVLNAGWNIKGVKNVDFVRTYDTVEFL 112
                   |:|| | : | :  ||| |  | :: |      :|:      :  || :
Db        334 GELALSATSGDITFNNNQVTNGSTSTRNAINIIDTAKV-TSIRAATG-QSIYFYDPITNP 391

Qy        113 SADTKTTTVNVESKD------------NGKK---TEVKIGAK-TSVIKE---------- 145
                | |:|:   |              :|:|   ||  | |  || |::
Db        392 GTAASTDTLNLNLADANSEIEYGGAIVFSGEKLSPTEKAIAANVTSTIRQPAVLARGDLV 451

Qy        146 -KDGKLVTGKDKGENGSS--TDEGEGLVTAKEV------------IDAVNKAGWRMKTT 189
               :||  || ||  ::  |    :|   ::|||             :|  |||   :||
Db        452 LRDGVTVTFKDLTQSPGSRILMDGGTTLSAKEANLSLNGLAVNLSSLDGTNKAA--LKTE 509

Qy        190 TAN 192
               |:
Db        510 AAD 512


RESULT 10
SLAP_LACAC
ID   SLAP_LACAC     STANDARD;      PRT;   444 AA.
AC   P35829;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   S-layer protein precursor (Surface layer protein) (SA-protein).
GN   SLPA.
OS   Lactobacillus acidophilus.
OC   Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OC   Lactobacillus.
OX   NCBI_TaxID=1579;
RN   [1]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 343-351 AND 440-444.
RC   STRAIN=ATCC 4356;
RX   MEDLINE=94012467; PubMed=8407780;
RA   Boot H.J., Kolen C.P.A.M., van Noort J.M., Pouwels P.H.;
RT   "S-layer protein of Lactobacillus acidophilus ATCC 4356:
RT   purification, expression in Escherichia coli, and nucleotide sequence
RT   of the corresponding gene.";
RL   J. Bacteriol. 175:6089-6096(1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 4356;
RX   MEDLINE=96099308; PubMed=8522531;
RA   Boot H.J., Kolen C.P.A.M., Pouwels P.H.;
RT   "Identification, cloning, and nucleotide sequence of a silent S-layer
RT   protein gene of Lactobacillus acidophilus ATCC 4356 which has
RT   extensive similarity with the S-layer protein gene of this species.";
RL   J. Bacteriol. 177:7222-7230(1995).
CC   -!- FUNCTION: THE S-LAYER IS A PARACRYSTALLINE MONO-LAYERED ASSEMBLY
CC       OF PROTEINS WHICH COAT THE SURFACE OF BACTERIA.
CC   -!- SUBCELLULAR LOCATION: CELL WALL. THIS BACTERIUM IS COVERED BY A
CC       S-LAYER WITH HEXAGONAL SYMMETRY.
CC   -!- SIMILARITY: SOME, TO THE S-LAYER PROTEIN OF L.BREVIS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X71412; CAA50535.1; -.
DR   EMBL; X89375; CAA61560.1; -.
DR   PIR; A36924; A36924.
DR   InterPro; IPR004903; SLAP.
DR   Pfam; PF03217; SLAP; 1.
KW   Signal; Glycoprotein; Cell wall; S-layer.
FT   SIGNAL        1     24       POTENTIAL.
FT   CHAIN        25    444       S-LAYER PROTEIN.
SQ   SEQUENCE   444 AA;  46570 MW;  2090732F89099161 CRC64;

  Query Match             9.6%;  Score 98;  DB 1;  Length 444;
  Best Local Similarity   22.9%;  Pred. No. 6.3;
  Matches   57;  Conservative   29;  Mismatches   87;  Indels   76;  Gaps   12;

Qy          7 AGDNLKIKQFTYSLKKDLTDLTSVGTEKLSFSANG----NKVNI-------TSDTKGLNF 55
              |  |:|   || :   : |: |:| | |:    ||      |:|       |:|   :||
Db        147 ANSNVK---FTGTNSDNQTE-TNVSTLKVKLDQNGVASLTNVSIANVYAINTTDNSNVNF 202

Qy         56 AKETAG---TNGDTTVHLNGIGST----------------------LTDRAASVKDVLNA 90
                 |:|   |||  :|: :  |                          |  | |: :| : |
Db        203 YDVTSGATVTNGAVSVNADNQGQVNVANVVAAINSKYFAAQYADKKLNTRTANTEDAIKA 262

Qy         91 GWNIK--GVKNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDG 148
                :   | :| :         | :  | |  |   |  |||   : :      : |
Db        263 ALKDQKIDVNSVGY--------FKAPHTFTVNVKATSNTNGKSATLPVVVTVPNVAEPTV 314

Qy        149 KLVTGK--------DKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQT----- 195
               |: :         ||      ||          :   |       ||| ||:|
Db        315 ASVSKRIMHNAYYYDKDAKRVGTDS----------VKRYNSVSVLPNTTTINGKTYYQVV 364

Qy        196 --GQA-DKF 201
                 |:| ||:
Db        365 ENGKAVDKY 373


RESULT 11
FLIC_SALMC
ID   FLIC_SALMC     STANDARD;      PRT;   504 AA.
AC   Q06981;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   Flagellin (Phase-1-D flagellin).
GN   FLIC.
OS   Salmonella moscow.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Salmonella.
OX   NCBI_TaxID=28146;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CDC;
RX   MEDLINE=93374829; PubMed=7690024;
RA   Masten B.J., Joys T.M.;
RT   "Molecular analyses of the Salmonella g. . . flagellar antigen
RT   complex.";
RL   J. Bacteriol. 175:5359-5365(1993).
CC   -!- FUNCTION: FLAGELLIN IS THE SUBUNIT PROTEIN WHICH POLYMERIZES TO
CC       FORM THE FILAMENTS OF BACTERIAL FLAGELLA.
CC   -!- MISCELLANEOUS: INDIVIDUAL SALMONELLA SEROTYPES USUALLY ALTERNATE
CC       BETWEEN THE PRODUCTION OF 2 ANTIGENIC FORMS OF FLAGELLA, TERMED
CC       PHASE-1 AND PHASE-2, EACH SPECIFIED BY SEPARATE STRUCTURAL GENES.
CC   -!- SIMILARITY: BELONGS TO THE BACTERIAL FLAGELLIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z15086; CAA78794.1; -.
DR   PIR; S33191; S33191.
DR   InterPro; IPR001029; Flagellin_C.
DR   InterPro; IPR001492; FlagellinN.
```

```
DR   Pfam; PF00700; Flagellin_C; 1.
DR   Pfam; PF00669; Flagellin_N; 1.
DR   PRINTS; PR00207; FLAGELLIN.
KW   Flagella.
FT   INIT_MET      0      0       BY SIMILARITY.
SQ   SEQUENCE   504 AA;  52791 MW;  FCAEA2180AF111A0 CRC64;

  Query Match             9.5%;  Score 97;  DB 1;  Length 504;
  Best Local Similarity   27.6%;  Pred. No. 8.5;
  Matches   47;  Conservative   22;  Mismatches   75;  Indels   26;  Gaps    8;

Qy         25 TDLTSVGT---EKLSFSANGNKVNITSDTKGLNFAKETA---GTNGDTTVHLNGIGSTLT 78
              |  :: ||   : :: :  | |   | | ||: |  :|    | ||  :  :||   |||
Db        249 TTKSTAGTAEAKAIAGAIKGGKEGDTFDYKGVTFTIDTKTGDGGNGKVSTTINGEKVTLT 308

Qy         79 DRAASVKDVLNAGWNIKGVKNVDFVRTYDTV---EFLSADTKTTTVNVESKD----NGKK 131
                | |:     |:                | :|   :| : | ||   : :  |    |  |
Db        309 -----VADIATGATNVNAATLQSSKNVYTSVVNGQF-TFDDKTKNESAKLSDLEANNAVK 362

Qy        132 TEVKI---GAKTSVIKEKDGKLVTGK----DKGENGSSTDEGEGLVTAKE 174
               | ||   ||: :     |   : ||    ||  :| ||   |    ||:
Db        363 GESKITVNGAEYTANATGDKITLAGKTMFIDKTASGVSTLINEDAAAAKK 412


RESULT 12
FLJB_SALTY
ID   FLJB_SALTY     STANDARD;      PRT;   505 AA.
AC   P52616; P97159;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Phase-2 flagellin.
GN   FLJB OR H2 OR STM2771.
OS   Salmonella typhimurium.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Salmonella.
OX   NCBI_TaxID=602;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SL 375;
RX   MEDLINE=95325331; PubMed=7541401;
RA   Vanegas R.A., Joys T.M.;
RT   "Molecular analyses of the phase-2 antigen complex 1,2,. . of
RT   Salmonella spp.";
RL   J. Bacteriol. 177:3863-3864(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX   MEDLINE=21534948; PubMed=11677609;
RA   McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA   Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA   Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA   Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA   Waterston R., Wilson R.K.;
RT   "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT   LT2.";
RL   Nature 413:852-856(2001).
RN   [3]
RP   SEQUENCE OF 1-37 FROM N.A.
RX   MEDLINE=82049491; PubMed=6271461;
RA   Silverman M., Zieg J., Mandel G., Simon M.;
RT   "Analysis of the functional components of the phase variation
RT   system.";
RL   Cold Spring Harb. Symp. Quant. Biol. 45:17-26(1981).
RN   [4]
RP   SEQUENCE OF 482-505 FROM N.A.
RC   STRAIN=SJ2353;
RA   Mingorance J., Tanaka S., Tominaga A., Enomoto M.;
RL   Submitted (JAN-1994) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: FLAGELLIN IS THE SUBUNIT PROTEIN WHICH POLYMERIZES TO
CC       FORM THE FILAMENTS OF BACTERIAL FLAGELLA.
CC   -!- MISCELLANEOUS: INDIVIDUAL SALMONELLA SEROTYPES USUALLY ALTERNATE
CC       BETWEEN THE PRODUCTION OF 2 ANTIGENIC FORMS OF FLAGELLA, TERMED
CC       PHASE-1 AND PHASE-2, EACH SPECIFIED BY SEPARATE STRUCTURAL GENES.
CC   -!- SIMILARITY: BELONGS TO THE BACTERIAL FLAGELLIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U17177; AAC43354.1; -.
DR   EMBL; AE008826; AAL21657.1; -.
DR   EMBL; V01370; CAA24655.1; -.
DR   EMBL; D26168; BAA05156.1; -.
DR   StyGene; SG10564; fljB.
DR   InterPro; IPR001029; Flagellin_C.
DR   InterPro; IPR001492; FlagellinN.
DR   Pfam; PF00700; Flagellin_C; 1.
DR   Pfam; PF00669; Flagellin_N; 1.
DR   PRINTS; PR00207; FLAGELLIN.
DR   ProDom; PD000316; Flagellin_C; 2.
KW   Flagella; Complete proteome.
FT   INIT_MET      0      0       BY SIMILARITY.
FT   CONFLICT     37     37       I -> S (IN REF. 3).
SQ   SEQUENCE   505 AA;  52404 MW;  3A5CC404AF7AF88B CRC64;

  Query Match             9.5%;  Score 96.5;  DB 1;  Length 505;
  Best Local Similarity   25.7%;  Pred. No. 9.2;
  Matches   49;  Conservative   25;  Mismatches   84;  Indels   33;  Gaps    8;

Qy         29 SVGTEKLSFSANGNKVNITSDTKGLNFAKETAGTNGDTTVHLNGIGSTLTDRAASVKDVL 88
              ||    : | |: ||  :|       |    |  |||  |::   | |:|  | : |  :
Db        219 SVTGGAVKFDADNNKYFVTIG----GFTGADAAKNGDYEVNVATDG-TVTLAAGATKTTM 273

Qy         89 NAGWNIKGVKNVDFVRTYDTVEFLSADTKTTTV--NVESKDNGKKTEVKIGAKTSVIKEK 146
               ||   |    :     ||   :||| |   :   |:: |      ||:        :|
Db        274 PAGATTK----TEVQELKDTPAVVSADAKNALIAGGVDATDANGAELVKMS-----YTDK 324

Qy        147 DGKLVTGKDKGENG-----SSTDEGEGLVTAKEVI---------DAVNKAG---WRMKTT 189
              :|| : |    : |     :  ||  | : ||             | |: |    : :
Db        325 NGKTIEGGYALKAGDKYYAADYDEATGAIKAKTTSYTAADGTTKTAANQLGGVDGKTEVV 384

Qy        190 TANGQTGQADK 200
              | :|:|  | |
Db        385 TIDGKTYNASK 395


RESULT 13
WAPA_BACSU
ID   WAPA_BACSU     STANDARD;      PRT;  2334 AA.
AC   Q07833;
DT   01-OCT-1994 (Rel. 30, Created)
DT   01-OCT-1994 (Rel. 30, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Wall-associated protein precursor.
GN   WAPA OR N17G.
OS   Bacillus subtilis.
OC   Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX   NCBI_TaxID=1423;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=168;
RX   MEDLINE=93302506; PubMed=8316082;
RA   Foster S.J.;
RT   "Molecular analysis of three major wall-associated proteins of
RT   Bacillus subtilis 168: evidence for processing of the product of a
RT   gene encoding a 258 kDa precursor two-domain ligand-binding
RT   protein.";
RL   Mol. Microbiol. 8:299-310(1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=168 / BGSC1A1;
```

```
RX   MEDLINE=95219088; PubMed=7704263;
RA   Yoshida K.-I., Sano H., Seki S., Oda M., Fujimura M., Fujita Y.;
RT   "Cloning and sequencing of a 29 kb region of the Bacillus subtilis
RT   genome containing the hut and wapA loci.";
RL   Microbiology 141:337-343(1995).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=168 / BGSC1A1;
RX   MEDLINE=97124196; PubMed=8969509;
RA   Yoshida K.-I., Shindo K., Sano H., Seki S., Fujimura M., Yanai N.,
RA   Miwa Y., Fujita Y.;
RT   "Sequencing of a 65 kb region of the Bacillus subtilis genome
RT   containing the lic and cel loci, and creation of a 177 kb contig
RT   covering the gnt-sacXY region.";
RL   Microbiology 142:3113-3123(1996).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=168;
RX   MEDLINE=98044033; PubMed=9384377;
RA   Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA   Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA   Borriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA   Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA   Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA   Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
RA   Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA   Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA   Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
RA   Guiseppi G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,
RA   Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA   Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA   Kobayashi Y., Koetter P., Koningstein G., Krogh S., Kumano M.,
RA   Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA   Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
RA   Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA   Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA   Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA   Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA   Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,
RA   Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA   Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
RA   Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA   Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA   Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA   Viari A., Wambutt R., Wedler E., Wedler H., Weitzenegger T.,
RA   Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA   Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;
RT   "The complete genome sequence of the Gram-positive bacterium Bacillus
RT   subtilis.";
RL   Nature 390:249-256(1997).
CC   -!- FUNCTION: STILL UNKNOWN. NOT INVOLVED IN CELL MEMBRANE METABOLISM,
CC       MOTILITY, SECRETION OR DIFFERENTIATION.
CC   -!- SUBCELLULAR LOCATION: MEMBRANE-ASSOCIATED. MAY BE RELEASED
CC       INTO THE MEDIUM.
CC   -!- DOMAIN: HAS TWO LIGAND-BINDING DOMAINS; THE N-TERMINUS, HAS THREE
CC       101 AA REPEATS WHICH ARE RESPONSIBLE FOR CELL WALL BINDING; THE
CC       C-TERMINUS CONSISTS OF TWO BLOCKS OF RESIDUES WITH A CONSERVED
CC       MOTIF REPEATED 31 TIMES.
CC   -!- SIMILARITY: THE REPEATED MOTIF IN THE C-TERMINUS HAS SOME
CC       SIMILARITY TO THE REPEAT IN E.COLI RHS GROUP OF PROTEINS (RHSA-D).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L05634; AAA22883.1; -.
DR   EMBL; D31856; BAA06656.1; -.
DR   EMBL; D29985; BAA06260.1; -.
DR   EMBL; D83026; BAA11683.1; -.
DR   EMBL; Z99124; CAB15959.1; -.
DR   PIR; S32920; S32920.
DR   SubtiList; BG10797; wapA.
DR   InterPro; IPR003305; CBM_CenC.
DR   InterPro; IPR006530; YD.
DR   Pfam; PF02018; CBM_4_9; 1.
DR   TIGRFAMs; TIGR01643; YD_repeat_2x; 17.
KW   Cell wall; Repeat; Signal; Complete proteome.
FT   SIGNAL       1     28       OR 32 (POTENTIAL).
FT   CHAIN       29   2334       WALL-ASSOCIATED PROTEIN.
FT   DOMAIN     504    869       3 X 101 AA APPROXIMATE TANDEM REPEATS.
FT   REPEAT     504    605       1-1.
FT   REPEAT     636    736       1-2.
FT   REPEAT     769    869       1-3.
FT   DOMAIN    1021   2139       31 X 21 AA APPROXIMATE TANDEM REPEATS OF
FT                               X(4)-G-X(4)-[YF]-X-D-X(2)-G-X(4).
FT   REPEAT    1021   1040       2-1.
FT   REPEAT    1042   1061       2-2.
FT   REPEAT    1063   1082       2-3.
FT   REPEAT    1083   1102       2-4.
FT   REPEAT    1109   1128       2-5.
FT   REPEAT    1129   1148       2-6.
FT   REPEAT    1150   1169       2-7.
FT   REPEAT    1174   1193       2-8.
FT   REPEAT    1199   1218       2-9.
FT   REPEAT    1219   1238       2-10.
FT   REPEAT    1646   1665       2-11.
FT   REPEAT    1667   1686       2-12.
FT   REPEAT    1690   1709       2-13.
FT   REPEAT    1711   1730       2-14.
FT   REPEAT    1732   1751       2-15.
FT   REPEAT    1753   1772       2-16.
FT   REPEAT    1795   1814       2-17.
FT   REPEAT    1820   1839       2-18.
FT   REPEAT    1840   1859       2-19.
FT   REPEAT    1861   1880       2-20.
FT   REPEAT    1887   1906       2-21.
FT   REPEAT    1908   1927       2-22.
FT   REPEAT    1929   1948       2-23.
FT   REPEAT    1969   1982       2-24 (APPROXIMATE).
FT   REPEAT    1983   2002       2-25.
FT   REPEAT    2008   2027       2-26.
FT   REPEAT    2028   2047       2-27.
FT   REPEAT    2051   2070       2-28.
FT   REPEAT    2071   2090       2-29.
FT   REPEAT    2093   2112       2-30.
FT   REPEAT    2120   2139       2-31.
SQ   SEQUENCE   2334 AA;  258329 MW;  B75138CCD278BAA3 CRC64;

  Query Match             9.5%;  Score 96.5;  DB 1;  Length 2334;
  Best Local Similarity   25.1%;  Pred. No. 52;
  Matches   64;  Conservative   24;  Mismatches   76;  Indels   91;  Gaps   16;

Qy         16 FTYSLKKDLTDLTSVG--------TEKLSFSANGNKVNITSDTKG--------LN-FAKE 58
              |:|  :::|     :|        ||   : ||||:  | |   |·       ||   ||
Db       1870 FSYD-ERNLVSSLHIGDKNGGDILTESYEYDANGNRTTINSSASGKVQYEYGKLNQLVKE 1928

Qy         59 TAGTNGDTTV---HLNGIGSTLTDRAASVKD----VLNAGWNI----------------K 95
                |: | ||     :| |:  |    ::||     :|| :||
Db       1929 ---THEDGTVIEYTYDGFGNRKT--VTTIKDGSSKTVNASFNIMNQLTKVNDESISYDKN 1983

Qy         96 GVKNVD--FVRTYDTVEFLSADTKT------TTVNVESKD-------NGKKTEVKI-GAK 139
              | :  |  |  |:|  : |:| ||        |   : |        ||| |     |
Db       1984 GNRTSDGKFTYTWDAEDNLTAVTKKGEDKPFATYKYDEKGNRIQKTVNGKVTNYFYDGDS 2043

Qy        140 TSVIKEKDG-----KLVTGKDKGENGSSTDEGE------------------GLVTAKEV 175
               :|: | |      |  |  | |:  | |: |:                  |   ||
Db       2044 LNVLYETDADNNVTKSYTYGDSGQLLSYTENGKKYFYHYNAHGDIIAISDSTGKTVAKYQ 2103

Qy        176 IDAVNKAGWRMKTTT 190
               ||     |   | |
Db       2104 YDA-----WGNPTKT 2113
```

```
RESULT 14
FLIC_SALMO
ID   FLIC_SALMO     STANDARD;      PRT;   504 AA.
AC   Q06973;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   Flagellin (Phase-1-C flagellin).
GN   FLIC.
OS   Salmonella montevideo.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Salmonella.
OX   NCBI_TaxID=115981;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 8387;
RX   MEDLINE=93374829; PubMed=7690024;
RA   Masten B.J., Joys T.M.;
RT   "Molecular analyses of the Salmonella g. . . flagellar antigen
RT   complex.";
RL   J. Bacteriol. 175:5359-5365(1993).
CC   -!- FUNCTION: FLAGELLIN IS THE SUBUNIT PROTEIN WHICH POLYMERIZES TO
CC       FORM THE FILAMENTS OF BACTERIAL FLAGELLA.
CC   -!- MISCELLANEOUS: INDIVIDUAL SALMONELLA SEROTYPES USUALLY ALTERNATE
CC       BETWEEN THE PRODUCTION OF 2 ANTIGENIC FORMS OF FLAGELLA, TERMED
CC       PHASE-1 AND PHASE-2, EACH SPECIFIED BY SEPARATE STRUCTURAL GENES.
CC   -!- SIMILARITY: BELONGS TO THE BACTERIAL FLAGELLIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z15069; CAA78778.1; -.
DR   PIR; S33190; S33190.
DR   InterPro; IPR001029; Flagellin_C.
DR   InterPro; IPR001492; FlagellinN.
DR   Pfam; PF00700; Flagellin_C; 1.
DR   Pfam; PF00669; Flagellin_N; 1.
DR   PRINTS; PR00207; FLAGELLIN.
KW   Flagella.
FT   INIT_MET      0      0       BY SIMILARITY.
SQ   SEQUENCE   504 AA;  52862 MW;  5C090577F21ECA67 CRC64;

  Query Match             9.4%;  Score 96;  DB 1;  Length 504;
  Best Local Similarity   30.2%;  Pred. No. 10;
  Matches   52;  Conservative   23;  Mismatches   67;  Indels   30;  Gaps   11;

Qy         25 TDLTSVGTEK---LSFSANGNKVNITSDTKGLNFAKETA---GTNGDTTVHLNGIGSTLT 78
              |  :: ||::   :: :  | |   | | ||: |  :|      ||  :  :||   |||
Db        249 TTKSAAGTDEAKAIAGAIKGGKEGDTFDYKGVTFTIDTKTGDDGNGKVSTTINGEKVTLT 308

Qy         79 --DRAASVKDVLNAGWNIKGVKNVDFVRTYDTV---EFLSADTKTTTVNVESKD----NG 129
                | |    || ||    ::  |||      | :|   :| : | ||   : :  |    |
Db        309 VADIATGATDV-NAA-TLQSSKNV-----YTSVVNGQF-TFDDKTKNESAKLSDLEANNA 360

Qy        130 KKTEVKI---GAKTSVIKEKDGKLVTGK----DKGENGSSTDEGEGLVTAKE 174
                | | ||   ||: :     |   : ||    ||  :| ||   |    ||:
Db        361 VKGESKITVNGAEYTANAAGDKVTLAGKTMFIDKTASGVSTLINEDAAAAKK 412


RESULT 15
RPOB_MYCGE
ID   RPOB_MYCGE     STANDARD;      PRT;  1390 AA.
AC   P47583; Q49439; Q49441;
DT   01-FEB-1996 (Rel. 33, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   DNA-directed RNA polymerase beta chain (EC 2.7.7.6) (Transcriptase
DE   beta chain) (RNA polymerase beta subunit).
GN   RPOB OR MG341.
OS   Mycoplasma genitalium.
OC   Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX   NCBI_TaxID=2097;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 33530 / G-37;
RX   MEDLINE=96026346; PubMed=7569993;
RA   Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,
RA   Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,
RA   Fritchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhrmann J.L.,
RA   Nguyen D.T., Utterback T.R., Saudek D.M., Phillips C.A., Merrick J.M.,
RA   Tomb J.-F., Dougherty B.A., Bott K.F., Hu P.-C., Lucier T.S.,
RA   Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;
RT   "The minimal gene complement of Mycoplasma genitalium.";
RL   Science 270:397-403(1995).
RN   [2]
RP   SEQUENCE OF 533-644 AND 945-1067 FROM N.A.
RC   STRAIN=ATCC 33530 / G-37;
RX   MEDLINE=94075230; PubMed=8253680;
RA   Peterson S.N., Hu P.-C., Bott K.F., Hutchison C.A. III;
RT   "A survey of the Mycoplasma genitalium genome by using random
RT   sequencing.";
RL   J. Bacteriol. 175:7918-7930(1993).
CC   -!- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION
CC       OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS
CC       SUBSTRATES.
CC   -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC       {RNA}(N).
CC   -!- SUBUNIT: THE ENZYME CONSISTS OF THE SIGMA CHAIN AND THE CORE
CC       ENZYME WHICH IS COMPOSED OF 2 ALPHA CHAINS, 1 BETA CHAIN, AND 1
CC       BETA' CHAIN.
CC   -!- SIMILARITY: Belongs to the RNA polymerase beta chain family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U39715; AAC71566.1; -.
DR   EMBL; U01737; AAD10547.1; -.
DR   EMBL; U01735; AAD10545.1; -.
DR   PIR; G64237; G64237.
DR   HSSP; Q9KWU7; 1HQM.
DR   TIGR; MG341; -.
DR   InterPro; IPR001572; RNA_pol_B.
DR   Pfam; PF04563; RNA_pol_Rpb2_1; 1.
DR   Pfam; PF04561; RNA_pol_Rpb2_2; 1.
DR   Pfam; PF04565; RNA_pol_Rpb2_3; 1.
DR   Pfam; PF00562; RNA_pol_Rpb2_6; 1.
DR   Pfam; PF04560; RNA_pol_Rpb2_7; 1.
DR   PROSITE; PS01166; RNA_POL_BETA; 1.
KW   Transferase; Transcription; DNA-directed RNA polymerase;
KW   Complete proteome.
FT   CONFLICT   1062   1062       H -> L (IN REF. 2).
SQ   SEQUENCE   1390 AA;  156330 MW;  4D1888A07680912C CRC64;

  Query Match             9.3%;  Score 95;  DB 1;  Length 1390;
  Best Local Similarity   25.0%;  Pred. No. 37;
  Matches   50;  Conservative   39;  Mismatches   65;  Indels   46;  Gaps   12;

Qy          5 LKAGDNLKIKQFTYSLKKDLTDLTSVG-TEKLSFSANGNKVNITSDTKGLNFAKETAG-- 61
              || |  | :|: |  ||:::  :      ::||    ||: :|:| | :: |||:
Db        413 LKDG-KLLLKKGTLLLKEEIDKIKQAAKNNEISFV---NKMQLTTDGKAVDLAKESLFYE 468

Qy         62 ------TNGDTTVHLNGIG----------STLTDRAASVKDVLNAGWNIKGVKNVDFVRT 105
                    || : :| :  ||           ||:|  ||:  |:|  : |
Db        469 TIDVYITNDNLSVSVPVIGIHNENDLNKAMTLSDFIASISYVINLPYGI---------GK 519
```

```
Qy        106 YDTVEFLSADTKTTTVN--VESKDNGKKTEVKIGAKTSVIKEK----DGKLVTGKDKGEN 159
              || :: |  : :   :|  : :|     | ::       :|||    || :  |:   |
Db        520 YDDIDHL-GNKRVKLINELITAKLESGFTRME-----RFLKEKLTIADG-VNRGQQINEE 572

Qy        160 GSSTDEGEGL-VTAKEVIDA 178
              |   ::||   :| | :|::
Db        573 GQVIEQGEKKELTIKSLINS 592


Search completed: October  6, 2003, 09:24:21
Job time : 4.10395 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: October 6, 2003, 09:13:10 ; Search time 14.7627 Seconds
(without alignments)
3513.485 Million cell updates/sec

Title: US-09-771-382-39
Perfect score: 1018
Sequence: 1 SANTLKAGDNLKIKQFTYSL.....AGWRMKTTTANGQTGQADKF 201

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_23:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	957.5	94.1	591	2	Q93QY3	Q93qy3 neisseria m
2	957.5	94.1	591	2	Q9JPS7	Q9jps7 neisseria m
3	957.5	94.1	591	16	Q9JR18	Q9jr18 neisseria m
4	953.5	93.7	592	2	Q9AQF0	Q9aqf0 neisseria m
5	947.5	93.1	590	2	Q9JPS3	Q9jps3 neisseria m
6	936.5	92.0	594	2	Q9JPI3	Q9jpi3 neisseria m
7	936.5	92.0	594	2	Q9JPS2	Q9jps2 neisseria m
8	931.5	91.5	595	2	Q9JPH0	Q9jph0 neisseria m
9	929.5	91.3	599	2	Q9JPS8	Q9jps8 neisseria m
10	928.5	91.2	600	2	Q9JPS6	Q9jps6 neisseria m
11	925.5	90.9	526	2	Q9JPS4	Q9jps4 neisseria m
12	925.5	90.9	598	2	Q93QY5	Q93qy5 neisseria m
13	925.5	90.9	598	2	Q9JPT0	Q9jpt0 neisseria m
14	925.5	90.9	599	2	Q9JPR8	Q9jpr8 neisseria m
15	923.5	90.7	530	2	Q9JPS1	Q9jps1 neisseria m
16	923.5	90.7	594	2	Q9JPH7	Q9jph7 neisseria m
17	923.5	90.7	598	2	Q9JPR7	Q9jpr7 neisseria m
18	922.5	90.6	594	2	Q93QY4	Q93qy4 neisseria m
19	921.5	90.5	598	2	Q9JPR9	Q9jpr9 neisseria m
20	921.5	90.5	598	2	Q9JPS0	Q9jps0 neisseria m
21	918.5	90.2	589	2	Q93QY1	Q93qy1 neisseria m
22	918.5	90.2	589	2	Q9JPI0	Q9jpi0 neisseria m
23	907.5	89.1	592	2	Q9JPS9	Q9jps9 neisseria m
24	907.5	89.1	592	2	Q93QY2	Q93qy2 neisseria m
25	903.5	88.8	592	16	Q9JQW4	Q9jqw4 neisseria m
26	895.5	88.0	600	2	Q9JPS5	Q9jps5 neisseria m
27	622.5	61.1	2353	2	P71401	P71401 haemophilus
28	618	60.7	1098	2	Q48152	Q48152 haemophilus
29	615.5	60.5	1096	2	Q8GM79	Q8gm79 haemophilus
30	595	58.4	1210	2	Q8GM75	Q8gm75 haemophilus
31	594	58.3	1210	2	Q8GM74	Q8gm74 haemophilus
32	570	56.0	1204	2	Q8GM76	Q8gm76 haemophilus
33	324	31.8	1002	2	Q8GM78	Q8gm78 haemophilus
34	322	31.6	1004	2	Q8GM77	Q8gm77 haemophilus
35	203.5	20.0	2314	2	Q8KQM8	Q8kqm8 moraxella c
36	199	19.5	1964	2	Q8KQM9	Q8kqm9 moraxella c
37	157.5	15.5	1264	2	Q8RQ61	Q8rq61 actinobacil
38	123.5	12.1	1633	16	Q8CMP4	Q8cmp4 staphylococ
39	123.5	12.1	1733	2	Q9KI14	Q9ki14 staphylococ
40	122.5	12.0	588	2	Q9S0T6	Q9s0t6 escherichia
41	122	12.0	2712	16	Q9F3X5	Q9f3x5 pasteurella
42	118	11.6	507	2	Q53821	Q53821 salmonella
43	118	11.6	507	2	Q54415	Q54415 salmonella
44	116.5	11.4	1557	2	Q9RNI2	Q9rni2 haemophilus
45	116	11.4	508	2	Q54515	Q54515 salmonella

ALIGNMENTS

RESULT 1
Q93QY3
ID Q93QY3 PRELIMINARY; PRT; 591 AA.
AC Q93QY3;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE NhhA outer membrane protein.
GN NHHA.
OS Neisseria meningitidis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=487;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EG329;
RA Peak I.R., Srikhanta Y., Dieckelman M., Moxon R., Jennings M.P.;
RT "Identification and characterization of a gene encoding a novel outer
RT membrane protein of Neisseria meningitidis.";
RL Submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF157606; AAK68867.1; -.
DR InterPro; IPR005594; YadA.
DR Pfam; PF03895; YadA; 1.
SQ SEQUENCE 591 AA; 62048 MW; C0DC600798859C65 CRC64;

Query Match 94.1%; Score 957.5; DB 2; Length 591;
Best Local Similarity 85.7%; Pred. No. 4.6e-48;
Matches 198; Conservative 0; Mismatches 0; Indels 33; Gaps 3;

```
Qy          4 TLKAGDNLKIKQ----FTYSLKKDLTDLTSVGTEKLSFSANGNKVNITSDTKGLNFAKET 59
              ||||||||||||    ||||||||||||||||||||||||||||||||||||||||||||
Db        109 TLKAGDNLKIKQNGTNFTYSLKKDLTDLTSVGTEKLSFSANGNKVNITSDTKGLNFAKET 168

Qy         60 AGTNGDTTVHLNGIGSTLTD----------------------RAASVKDVLNAGWNIKGV 97
              ||||||||||||||||||||                      ||||||||||||||||||
Db        169 AGTNGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGV 228

Qy         98 K-------NVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKL 150
              |       ||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        229 KPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKL 288

Qy        151 VTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKF 201
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        289 VTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKF 339


RESULT 2
Q9JPS7
ID   Q9JPS7      PRELIMINARY;      PRT;   591 AA.
AC   Q9JPS7;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Outer membrane protein GNA992.
GN   GNA992.
OS   Neisseria meningitidis.
OC   Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC   Neisseriaceae; Neisseria.
OX   NCBI_TaxID=487;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BZ147;
RX   MEDLINE=20175756; PubMed=10710308;
RA   Pizza M., Scarlato V., Masignani V., Giuliani M.M., Arico' B.,
RA   Comanducci M., Jennings G.T., Baldi L., Bartolini E., Capecchi B.,
RA   Galeotti C.L., Luzzi E., Manetti R., Marchetti E., Mora M., Nuti S.,
RA   Ratti G., Santini L., Savino S., Scarselli M., Storni E., Zuo P.,
RA   Broeker M., Hundt E., Knapp B., Blair E., Mason T., Tettelin H.,
RA   Hood D.W., Jeffries A.C., Saunders N.J., Granoff D.M., Venter C.,
RA   Moxon E.R., Grandi G., Rappuoli R.;
RT   "Identification of Vaccine Candidates Against Serogroup B
RT   Meningococcus by Whole-Genome Sequencing.";
RL   Science 287:1816-1820(2000).
DR   EMBL; AF226366; AAF42515.1; -.
DR   InterPro; IPR005594; YadA.
DR   Pfam; PF03895; YadA; 1.
SQ   SEQUENCE   591 AA;  62113 MW;  533453CAE5A91E1F CRC64;

  Query Match             94.1%;  Score 957.5;  DB 2;  Length 591;
  Best Local Similarity   85.7%;  Pred. No. 4.6e-48;
  Matches  198;  Conservative    0;  Mismatches    0;  Indels   33;  Gaps    3;

Qy          4 TLKAGDNLKIKQ----FTYSLKKDLTDLTSVGTEKLSFSANGNKVNITSDTKGLNFAKET 59
              ||||||||||||    ||||||||||||||||||||||||||||||||||||||||||
Db        109 TLKAGDNLKIKQNGTNFTYSLKKDLTDLTSVGTEKLSFSANGNKVNITSDTKGLNFAKET 168

Qy         60 AGTNGDTTVHLNGIGSTLTD----------------------RAASVKDVLNAGWNIKGV 97
              ||||||||||||||||||||                      ||||||||||||||||||
Db        169 AGTNGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGV 228

Qy         98 K-------NVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKL 150
              |       ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        229 KPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKL 288

Qy        151 VTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKF 201
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        289 VTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKF 339


RESULT 3
Q9JR18
ID   Q9JR18      PRELIMINARY;      PRT;   591 AA.
AC   Q9JR18;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Outer membrane protein GNA992 (Adhesin) (NhhA outer membrane
DE   protein).
GN   GNA992 OR NMB0992 OR NHHA.
OS   Neisseria meningitidis, and
OS   Neisseria meningitidis (serogroup B).
OC   Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC   Neisseriaceae; Neisseria.
OX   NCBI_TaxID=487, 491;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MC58 / Serogroup B, BZ169, BZ83, and H44/76;
RX   MEDLINE=20175756; PubMed=10710308;
RA   Pizza M., Scarlato V., Masignani V., Giuliani M.M., Arico' B.,
RA   Comanducci M., Jennings G.T., Baldi L., Bartolini E., Capecchi B.,
RA   Galeotti C.L., Luzzi E., Manetti R., Marchetti E., Mora M., Nuti S.,
RA   Ratti G., Santini L., Savino S., Scarselli M., Storni E., Zuo P.,
RA   Broeker M., Hundt E., Knapp B., Blair E., Mason T., Tettelin H.,
RA   Hood D.W., Jeffries A.C., Saunders N.J., Granoff D.M., Venter C.,
RA   Moxon E.R., Grandi G., Rappuoli R.;
RT   "Identification of Vaccine Candidates Against Serogroup B
RT   Meningococcus by Whole-Genome Sequencing.";
RL   Science 287:1816-1820(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MC58 / Serogroup B;
RX   MEDLINE=20175755; PubMed=10710307;
RA   Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
RA   Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
RA   Nelson W.C., Gwinn M.L., DeBoy R., Peterson J.D., Hickey E.K.,
RA   Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
RA   Mason T., Ciecko A., Parksey D.S., Blair E., Cittone H., Clark E.B.,
RA   Cotton M.D., Utterback T.R., Khouri H., Qin H., Vamathevan J.,
RA   Gill J., Scarlato V., Masignani V., Pizza M., Grandi G., Sun L.,
RA   Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;
RT   "Complete genome sequence of Neisseria meningitidis serogroup B strain
RT   MC58.";
RL   Science 287:1809-1815(2000).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   SPECIES=N.meningitidis; STRAIN=PMC21;
RA   Peak I.R., Srikhanta Y., Dieckelman M., Moxon R., Jennings M.P.;
RT   "Identification and characterization of a gene encoding a novel outer
RT   membrane protein of Neisseria meningitidis.";
RL   Submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF226375; AAF42524.1; -.
DR   EMBL; AE002450; AAF41395.1; -.
DR   EMBL; AF226367; AAF42516.1; -.
DR   EMBL; AF226370; AAF42519.1; -.
DR   EMBL; AF226374; AAF42523.1; -.
DR   EMBL; AF157611; AAK68872.1; -.
DR   TIGR; NMB0992; -.
DR   InterPro; IPR005594; YadA.
DR   Pfam; PF03895; YadA; 1.
KW   Complete proteome.
SQ   SEQUENCE   591 AA;  62112 MW;  7C22F3CAE7F73EC6 CRC64;

  Query Match             94.1%;  Score 957.5;  DB 16;  Length 591;
  Best Local Similarity   85.7%;  Pred. No. 4.6e-48;
  Matches  198;  Conservative    0;  Mismatches    0;  Indels   33;  Gaps    3;

Qy          4 TLKAGDNLKIKQ----FTYSLKKDLTDLTSVGTEKLSFSANGNKVNITSDTKGLNFAKET 59
              ||||||||||||    ||||||||||||||||||||||||||||||||||||||||||
Db        109 TLKAGDNLKIKQNGTNFTYSLKKDLTDLTSVGTEKLSFSANGNKVNITSDTKGLNFAKET 168

Qy         60 AGTNGDTTVHLNGIGSTLTD----------------------RAASVKDVLNAGWNIKGV 97
              ||||||||||||||||||||                      ||||||||||||||||||
Db        169 AGTNGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGV 228

Qy         98 K-------NVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKL 150
              |       ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        229 KPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKL 288

Qy        151 VTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKF 201
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        289 VTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKF 339


RESULT 4
Q9AQF0
```

```
ID   Q9AQF0      PRELIMINARY;      PRT;   592 AA.
AC   Q9AQF0;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Outer membrane protein.
GN   NHHA.
OS   Neisseria meningitidis.
OC   Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC   Neisseriaceae; Neisseria.
OX   NCBI_TaxID=487;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MC58;
RA   Peak I.R., Srikhanta Y., Dieckelman M., Moxon E.R., Jennings M.P.;
RT   "Identification and characterisation of a gene encoding a novel outer
RT   membrane protein of Neisseria meningitidis.";
RL   Submitted (FEB-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF125375; AAK09243.1; -.
DR   InterPro; IPR005594; YadA.
DR   Pfam; PF03895; YadA; 1.
SQ   SEQUENCE   592 AA;  62290 MW;  168986A97381EFC5 CRC64;

  Query Match             93.7%;  Score 953.5;  DB 2;  Length 592;
  Best Local Similarity   85.3%;  Pred. No. 7.9e-48;
  Matches  197;  Conservative    0;  Mismatches    1;  Indels   33;  Gaps    3;

Qy          4 TLKAGDNLKIKQ----FTYSLKKDLTDLTSVGTEKLSFSANGNKVNITSDTKGLNFAKET 59
              ||||||||||||    ||||||||||||||||||||||||||||||||||||||||||||
Db        110 TLKAGDNLKIKQNGTNFTYSLKKDLTDLTSVGTEKLSFSANGNKVNITSDTKGLNFAKET 169

Qy         60 AGTNGDTTVHLNGIGSTLTD----------------------RAASVKDVLNAGWNIKGV 97
              ||||||||||||||||||||                      ||||||||||||||||||
Db        170 AGTNGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGV 229

Qy         98 K-------NVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKL 150
              |       ||||||||||||||||||||||||||||||||||| ||||||||||||
Db        230 KPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGVKTSVIKEKDGKL 289

Qy        151 VTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKF 201
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        290 VTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKF 340


RESULT 5
Q9JPS3
ID   Q9JPS3      PRELIMINARY;      PRT;   590 AA.
AC   Q9JPS3;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Outer membrane protein GNA992.
GN   GNA992.
OS   Neisseria meningitidis.
OC   Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC   Neisseriaceae; Neisseria.
OX   NCBI_TaxID=487;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NGE28;
RX   MEDLINE=20175756; PubMed=10710308;
RA   Pizza M., Scarlato V., Masignani V., Giuliani M.M., Arico' B.,
RA   Comanducci M., Jennings G.T., Baldi L., Bartolini E., Capecchi B.,
RA   Galeotti C.L., Luzzi E., Manetti R., Marchetti E., Mora M., Nuti S.,
RA   Ratti G., Santini L., Savino S., Scarselli M., Storni E., Zuo P.,
RA   Broeker M., Hundt E., Knapp B., Blair E., Mason T., Tettelin H.,
RA   Hood D.W., Jeffries A.C., Saunders N.J., Granoff D.M., Venter C.,
RA   Moxon E.R., Grandi G., Rappuoli R.;
RT   "Identification of Vaccine Candidates Against Serogroup B
RT   Meningococcus by Whole-Genome Sequencing.";
RL   Science 287:1816-1820(2000).
DR   EMBL; AF226378; AAF42527.1; -.
DR   InterPro; IPR005594; YadA.
DR   Pfam; PF03895; YadA; 1.
SQ   SEQUENCE   590 AA;  61661 MW;  8AA476AC300D80C8 CRC64;

  Query Match             93.1%;  Score 947.5;  DB 2;  Length 590;
  Best Local Similarity   84.8%;  Pred. No. 1.8e-47;
  Matches  196;  Conservative    1;  Mismatches    1;  Indels   33;  Gaps    3;

Qy          4 TLKAGDNLKIKQ----FTYSLKKDLTDLTSVGTEKLSFSANGNKVNITSDTKGLNFAKET 59
              ||||||||||||    ||||||||||||||||||||||||||||||||||||||||||||
Db        107 TLKAGDNLKIKQNGTNFTYSLKKDLTDLTSVGTEKLSFSANGNKVNITSDTKGLNFAKET 166

Qy         60 AGTNGDTTVHLNGIGSTLTD----------------------RAASVKDVLNAGWNIKGV 97
              ||||||||||||||||||||                      ||||||||||||||||||
Db        167 AGTNGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAGWNIKGV 226

Qy         98 K-------NVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIKEKDGKL 150
              |       ||||||||||||||||||||||||||||||||||:||||||||||||||||||
Db        227 KPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKRTEVKIGAKTSVIKEKDGKL 286

Qy        151 VTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKF 201
              |||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        287 VTGKGKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKF 337


RESULT 6
Q9JPI3
ID   Q9JPI3      PRELIMINARY;      PRT;   594 AA.
AC   Q9JPI3;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Outer membrane protein GNA992.
GN   GNA992.
OS   Neisseria meningitidis.
OC   Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC   Neisseriaceae; Neisseria.
OX   NCBI_TaxID=487;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NG3/88, and BZ232;
RX   MEDLINE=20175756; PubMed=10710308;
RA   Pizza M., Scarlato V., Masignani V., Giuliani M.M., Arico' B.,
RA   Comanducci M., Jennings G.T., Baldi L., Bartolini E., Capecchi B.,
RA   Galeotti C.L., Luzzi E., Manetti R., Marchetti E., Mora M., Nuti S.,
RA   Ratti G., Santini L., Savino S., Scarselli M., Storni E., Zuo P.,
RA   Broeker M., Hundt E., Knapp B., Blair E., Mason T., Tettelin H.,
RA   Hood D.W., Jeffries A.C., Saunders N.J., Granoff D.M., Venter C.,
RA   Moxon E.R., Grandi G., Rappuoli R.;
RT   "Identification of Vaccine Candidates Against Serogroup B
RT   Meningococcus by Whole-Genome Sequencing.";
RL   Science 287:1816-1820(2000).
DR   EMBL; AF226376; AAF42525.1; -.
DR   EMBL; AF226369; AAF42518.1; -.
DR   InterPro; IPR005594; YadA.
DR   Pfam; PF03895; YadA; 1.
SQ   SEQUENCE   594 AA;  62086 MW;  1B25E03B90D04B46 CRC64;

  Query Match             92.0%;  Score 936.5;  DB 2;  Length 594;
  Best Local Similarity   82.3%;  Pred. No. 7.7e-47;
  Matches  195;  Conservative    0;  Mismatches    3;  Indels   39;  Gaps    3;

Qy          4 TLKAGDNLKIKQ----------FTYSLKKDLTDLTSVGTEKLSFSANGNKVNITSDTKGL 53
              ||||||||||||          ||||||||||||||| |||||| ||| ||||||||||
Db        105 TLKAGDNLKIKQNTDENTNASSFTYSLKKDLTDLTSVETEKLSFGANGKKVNITSDTKGL 164

Qy         54 NFAKETAGTNGDTTVHLNGIGSTLTD----------------------RAASVKDVLNAG 91
              |||||||||||||||||||||||||                      ||||||||||
Db        165 NFAKETAGTNGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAG 224

Qy         92 WNIKGVK-------NVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIK 144
              |||||||       ||||||||||||||||||||||||||||||||||||||||||||||
Db        225 WNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIK 284
```

```
Qy        145 EKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKF 201
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        285 EKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKF 341


RESULT 7
Q9JPS2
ID   Q9JPS2      PRELIMINARY;      PRT;   594 AA.
AC   Q9JPS2;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Outer membrane protein GNA992.
GN   GNA992.
OS   Neisseria meningitidis.
OC   Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC   Neisseriaceae; Neisseria.
OX   NCBI_TaxID=487;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NGE31;
RX   MEDLINE=20175756; PubMed=10710308;
RA   Pizza M., Scarlato V., Masignani V., Giuliani M.M., Arico' B.,
RA   Comanducci M., Jennings G.T., Baldi L., Bartolini E., Capecchi B.,
RA   Galeotti C.L., Luzzi E., Manetti R., Marchetti E., Mora M., Nuti S.,
RA   Ratti G., Santini L., Savino S., Scarselli M., Storni E., Zuo P.,
RA   Broeker M., Hundt E., Knapp B., Blair E., Mason T., Tettelin H.,
RA   Hood D.W., Jeffries A.C., Saunders N.J., Granoff D.M., Venter C.,
RA   Moxon E.R., Grandi G., Rappuoli R.;
RT   "Identification of Vaccine Candidates Against Serogroup B
RT   Meningococcus by Whole-Genome Sequencing.";
RL   Science 287:1816-1820(2000).
DR   EMBL; AF226379; AAF42528.1; -.
DR   InterPro; IPR005594; YadA.
DR   Pfam; PF03895; YadA; 1.
SQ   SEQUENCE   594 AA;  62114 MW;  1E2A63A78F53D256 CRC64;

  Query Match             92.0%;  Score 936.5;  DB 2;  Length 594;
  Best Local Similarity   82.3%;  Pred. No. 7.7e-47;
  Matches  195;  Conservative    0;  Mismatches    3;  Indels   39;  Gaps    3;

Qy          4 TLKAGDNLKIKQ----------FTYSLKKDLTDLTSVGTEKLSFSANGNKVNITSDTKGL 53
              ||||||||||||          ||||||||||||||| |||||| ||| ||||||||||||
Db        105 TLKAGDNLKIKQNTDENTNASSFTYSLKKDLTDLTSVETEKLSFGANGKKVNITSDTKGL 164

Qy         54 NFAKETAGTNGDTTVHLNGIGSTLTD----------------------RAASVKDVLNAG 91
              |||||||||||||||||||||||||                      ||||||||||||
Db        165 NFAKETAGTNGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAG 224

Qy         92 WNIKGVK-------NVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIK 144
              |||||||       ||||||||||||||||||||||||||||||||||||||||||||||
Db        225 WNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIK 284

Qy        145 EKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKF 201
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        285 EKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKF 341


RESULT 8
Q9JPH0
ID   Q9JPH0      PRELIMINARY;      PRT;   595 AA.
AC   Q9JPH0;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Outer membrane protein GNA992.
GN   GNA992.
OS   Neisseria meningitidis.
OC   Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC   Neisseriaceae; Neisseria.
OX   NCBI_TaxID=487;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=528, and 1000;
RX   MEDLINE=20175756; PubMed=10710308;
RA   Pizza M., Scarlato V., Masignani V., Giuliani M.M., Arico' B.,
RA   Comanducci M., Jennings G.T., Baldi L., Bartolini E., Capecchi B.,
RA   Galeotti C.L., Luzzi E., Manetti R., Marchetti E., Mora M., Nuti S.,
RA   Ratti G., Santini L., Savino S., Scarselli M., Storni E., Zuo P.,
RA   Broeker M., Hundt E., Knapp B., Blair E., Mason T., Tettelin H.,
RA   Hood D.W., Jeffries A.C., Saunders N.J., Granoff D.M., Venter C.,
RA   Moxon E.R., Grandi G., Rappuoli R.;
RT   "Identification of Vaccine Candidates Against Serogroup B
RT   Meningococcus by Whole-Genome Sequencing.";
RL   Science 287:1816-1820(2000).
DR   EMBL; AF226360; AAF42509.1; -.
DR   EMBL; AF226356; AAF42505.1; -.
DR   InterPro; IPR005594; YadA.
DR   Pfam; PF03895; YadA; 1.
SQ   SEQUENCE   595 AA;  62120 MW;  8212C96380142BFC CRC64;

  Query Match             91.5%;  Score 931.5;  DB 2;  Length 595;
  Best Local Similarity   81.0%;  Pred. No. 1.5e-46;
  Matches  192;  Conservative    4;  Mismatches    2;  Indels   39;  Gaps    3;

Qy          4 TLKAGDNLKIKQ----------FTYSLKKDLTDLTSVGTEKLSFSANGNKVNITSDTKGL 53
              ||||||||||||          |||||||||||||||||:||| ||||||||||||||||
Db        106 TLKAGDNLKIKQNTDENTNASSFTYSLKKDLTDLTSVGTEELSFGANGNKVNITSDTKGL 165

Qy         54 NFAKETAGTNGDTTVHLNGIGSTLTD----------------------RAASVKDVLNAGWN 93
              ||||:||||||||||||||||||||                      ||||:|||||||||
Db        166 NFAKKTAGTNGDTTVHLNGIGSTLTDTLAGSSASHVDAGNQSTHYTRAASIKDVLNAGWN 225

Qy         94 IKGVK---------NVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIK 144
              |||||         ||||||||||||||||||||||||||||||||:|||||||||||||
Db        226 IKGVKTGSTTGQSENVDFVRTYDTVEFLSADTKTTTVNVESKDNGKRTEVKIGAKTSVIK 285

Qy        145 EKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKF 201
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db        286 EKDGKLVTGKGKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKF 342


RESULT 9
Q9JPS8
ID   Q9JPS8      PRELIMINARY;      PRT;   599 AA.
AC   Q9JPS8;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Outer membrane protein GNA992.
GN   GNA992.
OS   Neisseria meningitidis.
OC   Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC   Neisseriaceae; Neisseria.
OX   NCBI_TaxID=487;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A22;
RX   MEDLINE=20175756; PubMed=10710308;
RA   Pizza M., Scarlato V., Masignani V., Giuliani M.M., Arico' B.,
RA   Comanducci M., Jennings G.T., Baldi L., Bartolini E., Capecchi B.,
RA   Galeotti C.L., Luzzi E., Manetti R., Marchetti E., Mora M., Nuti S.,
RA   Ratti G., Santini L., Savino S., Scarselli M., Storni E., Zuo P.,
RA   Broeker M., Hundt E., Knapp B., Blair E., Mason T., Tettelin H.,
RA   Hood D.W., Jeffries A.C., Saunders N.J., Granoff D.M., Venter C.,
RA   Moxon E.R., Grandi G., Rappuoli R.;
RT   "Identification of Vaccine Candidates Against Serogroup B
RT   Meningococcus by Whole-Genome Sequencing.";
RL   Science 287:1816-1820(2000).
DR   EMBL; AF226364; AAF42513.1; -.
DR   InterPro; IPR005594; YadA.
DR   Pfam; PF03895; YadA; 1.
SQ   SEQUENCE   599 AA;  62693 MW;  46C2E974AF7F78E9 CRC64;

  Query Match             91.3%;  Score 929.5;  DB 2;  Length 599;
  Best Local Similarity   80.1%;  Pred. No. 2e-46;
```

```
  Matches  193;  Conservative    2;  Mismatches    3;  Indels   43;  Gaps    3;

Qy          4 TLKAGDNLKIKQ--------------FTYSLKKDLTDLTSVGTEKLSFSANGNKVNITSD 49
              ||||||||||||              ||||||||||||||| |||||| ||||||||||
Db        106 TLKAGDNLKIKQNTDENTDENTNASSFTYSLKKDLTDLTSVETEKLSFGANGNKVNITSD 165

Qy         50 TKGLNFAKETAGTNGDTTVHLNGIGSTLTD--------------------RAASVKDVLN 89
              ||||||||||||||||||||||||||||||                    ||||:|||||
Db        166 TKGLNFAKETAGTNGDTTVHLNGIGSTLTDTLAGSSASHVDAGNQSTHYTRAASIKDVLN 225

Qy         90 AGWNIKGVK---------NVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKT 140
              |||||||||         ||||||||||||||||||||||||||||||:|||||||||
Db        226 AGWNIKGVKTGSTTGQSENVDFVRTYDTVEFLSADTKTTTVNVESKDNGKRTEVKIGAKT 285

Qy        141 SVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADK 200
              |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        286 SVIKEKDGKLVTGKGKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADK 345

Qy        201 F 201
              |
Db        346 F 346


RESULT 10
Q9JPS6
ID   Q9JPS6      PRELIMINARY;      PRT;   600 AA.
AC   Q9JPS6;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Outer membrane protein GNA992.
GN   GNA992.
OS   Neisseria meningitidis.
OC   Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC   Neisseriaceae; Neisseria.
OX   NCBI_TaxID=487;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=E26;
RX   MEDLINE=20175756; PubMed=10710308;
RA   Pizza M., Scarlato V., Masignani V., Giuliani M.M., Arico' B.,
RA   Comanducci M., Jennings G.T., Baldi L., Bartolini E., Capecchi B.,
RA   Galeotti C.L., Luzzi E., Manetti R., Marchetti E., Mora M., Nuti S.,
RA   Ratti G., Santini L., Savino S., Scarselli M., Storni E., Zuo P.,
RA   Broeker M., Hundt E., Knapp B., Blair E., Mason T., Tettelin H.,
RA   Hood D.W., Jeffries A.C., Saunders N.J., Granoff D.M., Venter C.,
RA   Moxon E.R., Grandi G., Rappuoli R.;
RT   "Identification of Vaccine Candidates Against Serogroup B
RT   Meningococcus by Whole-Genome Sequencing.";
RL   Science 287:1816-1820(2000).
DR   EMBL; AF226371; AAF42520.1; -.
DR   InterPro; IPR005594; YadA.
DR   Pfam; PF03895; YadA; 1.
SQ   SEQUENCE   600 AA;  62762 MW;  36256963E0598CD1 CRC64;

  Query Match             91.2%;  Score 928.5;  DB 2;  Length 600;
  Best Local Similarity   81.4%;  Pred. No. 2.3e-46;
  Matches  193;  Conservative    1;  Mismatches    4;  Indels   39;  Gaps    3;

Qy          4 TLKAGDNLKIKQ----------FTYSLKKDLTDLTSVGTEKLSFSANGNKVNITSDTKGL 53
              ||||||||||||          |||||||:||||||||||||| ||||||||||||||
Db        112 TLKAGDNLKIKQNTDENTNASSFTYSLKKELTDLTSVGTEKLSFGANGNKVNITSDTKGL 171

Qy         54 NFAKETAGTNGDTTVHLNGIGSTLTD----------------------RAASVKDVLNAG 91
              |||||||||||||||||||||||||                      ||||||||||||
Db        172 NFAKETAGTNGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVKDVLNAG 231

Qy         92 WNIKGVK-------NVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIK 144
              |||||||       ||||| ||||||||||||||||||||||||||||||||||||||||
Db        232 WNIKGVKPGTTASDNVDFVHTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKTSVIK 291

Qy        145 EKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADKF 201
              |||||||||| |||||||||||||||||||||||||||||||||||||||||| ||||||
Db        292 EKDGKLVTGKGKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQKGQADKF 348


RESULT 11
Q9JPS4
ID   Q9JPS4      PRELIMINARY;      PRT;   526 AA.
AC   Q9JPS4;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Outer membrane protein GNA992.
GN   GNA992.
OS   Neisseria meningitidis.
OC   Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC   Neisseriaceae; Neisseria.
OX   NCBI_TaxID=487;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NG6/88;
RX   MEDLINE=20175756; PubMed=10710308;
RA   Pizza M., Scarlato V., Masignani V., Giuliani M.M., Arico' B.,
RA   Comanducci M., Jennings G.T., Baldi L., Bartolini E., Capecchi B.,
RA   Galeotti C.L., Luzzi E., Manetti R., Marchetti E., Mora M., Nuti S.,
RA   Ratti G., Santini L., Savino S., Scarselli M., Storni E., Zuo P.,
RA   Broeker M., Hundt E., Knapp B., Blair E., Mason T., Tettelin H.,
RA   Hood D.W., Jeffries A.C., Saunders N.J., Granoff D.M., Venter C.,
RA   Moxon E.R., Grandi G., Rappuoli R.;
RT   "Identification of Vaccine Candidates Against Serogroup B
RT   Meningococcus by Whole-Genome Sequencing.";
RL   Science 287:1816-1820(2000).
DR   EMBL; AF226377; AAF42526.1; -.
DR   InterPro; IPR005594; YadA.
DR   Pfam; PF03895; YadA; 1.
SQ   SEQUENCE   526 AA;  54732 MW;  F543C4234AF807CA CRC64;

  Query Match             90.9%;  Score 925.5;  DB 2;  Length 526;
  Best Local Similarity   80.1%;  Pred. No. 2.9e-46;
  Matches  193;  Conservative    1;  Mismatches    4;  Indels   43;  Gaps    3;

Qy          4 TLKAGDNLKIKQ--------------FTYSLKKDLTDLTSVGTEKLSFSANGNKVNITSD 49
              ||||||||||||              ||||||||||||||| |||||| ||||||||||
Db     .    34 TLKAGDNLKIKQNTNKNTNENTNDSSFTYSLKKDLTDLTSVETEKLSFGANGNKVNITSD 93

Qy         50 TKGLNFAKETAGTNGDTTVHLNGIGSTLTD----------------------RAASVKDV 87
              |||||||||||||||||||||||||||||                      ||||||||
Db         94 TKGLNFAKETAGTNGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDKKKRAASVKDV 153

Qy         88 LNAGWNIKGVK-------NVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKT 140
              ||||||||||         ||||| ||||||||||||||||||||||||:|||||||||
Db        154 LNAGWNIKGVKPGTTASDNVDFVHTYDTVEFLSADTKTTTVNVESKDNGKRTEVKIGAKT 213

Qy        141 SVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADK 200
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        214 SVIKEKDGKLVTGKGKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADK 273

Qy        201 F 201
              |
Db        274 F 274


RESULT 12
Q93QY5
ID   Q93QY5      PRELIMINARY;      PRT;   598 AA.
AC   Q93QY5;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   NhhA outer membrane protein.
GN   NHHA.
OS   Neisseria meningitidis.
OC   Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC   Neisseriaceae; Neisseria.
OX   NCBI_TaxID=487;
```

```
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BZ10;
RA   Peak I.R., Srikhanta Y., Dieckelman M., Moxon R., Jennings M.P.;
RT   "Identification and characterization of a gene encoding a novel outer
RT   membrane protein of Neisseria meningitidis.";
RL   Submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF157603; AAK68864.1; -.
DR   InterPro; IPR005594; YadA.
DR   Pfam; PF03895; YadA; 1.
SQ   SEQUENCE   598 AA;  62687 MW;  18CEFFE6410A15DF CRC64;

  Query Match             90.9%;  Score 925.5;  DB 2;  Length 598;
  Best Local Similarity   79.8%;  Pred. No. 3.4e-46;
  Matches  194;  Conservative    1;  Mismatches    5;  Indels   43;  Gaps    3;

Qy          2 ANTLKAGDNLKIKQ--------------FTYSLKKDLTDLTSVGTEKLSFSANGNKVNIT 47
              | ||||||||||||              |||||||||||||| |||||| |||||||||
Db        103 AITLKAGDNLKIKQNTNENTNENTNDSSFTYSLKKDLTDLTSVETEKLSFGANGNKVNIT 162

Qy         48 SDTKGLNFAKETAGTNGDTTVHLNGIGSTLTD----------------------RAASVK 85
              |||||||||||||||| ||||||||||||                        ||||||
Db        163 SDTKGLNFAKETAGTNGDPTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVK 222

Qy         86 DVLNAGWNIKGVK-------NVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGA 138
              |||||||||||||       |||||||||||||||||||||||||||||:|||||||
Db        223 DVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKRTEVKIGA 282

Qy        139 KTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQA 198
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        283 KTSVIKEKDGKLVTGKGKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQA 342

Qy        199 DKF 201
              |||
Db        343 DKF 345


RESULT 13
Q9JPT0
ID   Q9JPT0      PRELIMINARY;      PRT;   598 AA.
AC   Q9JPT0;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Outer membrane protein GNA992.
GN   GNA992.
OS   Neisseria meningitidis.
OC   Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC   Neisseriaceae; Neisseria.
OX   NCBI_TaxID=487;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=2996;
RX   MEDLINE=20175756; PubMed=10710308;
RA   Pizza M., Scarlato V., Masignani V., Giuliani M.M., Arico' B.,
RA   Comanducci M., Jennings G.T., Baldi L., Bartolini E., Capecchi B.,
RA   Galeotti C.L., Luzzi E., Manetti R., Marchetti E., Mora M., Nuti S.,
RA   Ratti G., Santini L., Savino S., Scarselli M., Storni E., Zuo P.,
RA   Broeker M., Hundt E., Knapp B., Blair E., Mason T., Tettelin H.,
RA   Hood D.W., Jeffries A.C., Saunders N.J., Granoff D.M., Venter C.,
RA   Moxon E.R., Grandi G., Rappuoli R.;
RT   "Identification of Vaccine Candidates Against Serogroup B
RT   Meningococcus by Whole-Genome Sequencing.";
RL   Science 287:1816-1820(2000).
DR   EMBL; AF226359; AAF42508.1; -.
DR   InterPro; IPR005594; YadA.
DR   Pfam; PF03895; YadA; 1.
SQ   SEQUENCE   598 AA;  62763 MW;  63A6A3BD7F0F2EE3 CRC64;

  Query Match             90.9%;  Score 925.5;  DB 2;  Length 598;
  Best Local Similarity   79.8%;  Pred. No. 3.4e-46;
  Matches  194;  Conservative    1;  Mismatches    5;  Indels   43;  Gaps    3;

Qy          2 ANTLKAGDNLKIKQ--------------FTYSLKKDLTDLTSVGTEKLSFSANGNKVNIT 47
              | ||||||||||||              |||||||||||||| |||||| |||||||||
Db        103 AITLKAGDNLKIKQNTNENTNENTNDSSFTYSLKKDLTDLTSVETEKLSFGANGNKVNIT 162

Qy         48 SDTKGLNFAKETAGTNGDTTVHLNGIGSTLTD----------------------RAASVK 85
              |||||||||||||||| ||||||||||||                        ||||||
Db        163 SDTKGLNFAKETAGTNGDPTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDEKKRAASVK 222

Qy         86 DVLNAGWNIKGVK-------NVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGA 138
              |||||||||||||       |||||||||||||||||||||||||||||:|||||||
Db        223 DVLNAGWNIKGVKPGTTASDNVDFVRTYDTVEFLSADTKTTTVNVESKDNGKRTEVKIGA 282

Qy        139 KTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQA 198
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        283 KTSVIKEKDGKLVTGKGKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQA 342

Qy        199 DKF 201
              |||
Db        343 DKF 345


RESULT 14
Q9JPR8
ID   Q9JPR8      PRELIMINARY;      PRT;   599 AA.
AC   Q9JPR8;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Outer membrane protein GNA992 (NhhA outer membrane protein).
GN   GNA992 OR NHHA.
OS   Neisseria meningitidis.
OC   Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC   Neisseriaceae; Neisseria.
OX   NCBI_TaxID=487;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NGH38;
RX   MEDLINE=20175756; PubMed=10710308;
RA   Pizza M., Scarlato V., Masignani V., Giuliani M.M., Arico' B.,
RA   Comanducci M., Jennings G.T., Baldi L., Bartolini E., Capecchi B.,
RA   Galeotti C.L., Luzzi E., Manetti R., Marchetti E., Mora M., Nuti S.,
RA   Ratti G., Santini L., Savino S., Scarselli M., Storni E., Zuo P.,
RA   Broeker M., Hundt E., Knapp B., Blair E., Mason T., Tettelin H.,
RA   Hood D.W., Jeffries A.C., Saunders N.J., Granoff D.M., Venter C.,
RA   Moxon E.R., Grandi G., Rappuoli R.;
RT   "Identification of Vaccine Candidates Against Serogroup B
RT   Meningococcus by Whole-Genome Sequencing.";
RL   Science 287:1816-1820(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=H38;
RA   Peak I.R., Srikhanta Y., Dieckelman M., Moxon R., Jennings M.P.;
RT   "Identification and characterization of a gene encoding a novel outer
RT   membrane protein of Neisseria meningitidis.";
RL   Submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF226383; AAF42532.1; -.
DR   EMBL; AF157608; AAK68869.1; -.
DR   InterPro; IPR005594; YadA.
DR   Pfam; PF03895; YadA; 1.
SQ   SEQUENCE   599 AA;  62844 MW;  BBA16EBF53C1970C CRC64;

  Query Match             90.9%;  Score 925.5;  DB 2;  Length 599;
  Best Local Similarity   80.1%;  Pred. No. 3.4e-46;
  Matches  193;  Conservative    1;  Mismatches    4;  Indels   43;  Gaps    3;

Qy          4 TLKAGDNLKIKQ--------------FTYSLKKDLTDLTSVGTEKLSFSANGNKVNITSD 49
              ||||||||||||              |||||||||||||| |||||| ||||||||||
Db        106 TLKAGDNLKIKQNTNKNTNENTNDSSFTYSLKKDLTDLTSVETEKLSFGANGNKVNITSD 165

Qy         50 TKGLNFAKETAGTNGDTTVHLNGIGSTLTD----------------------RAASVKDV 87
              ||||||||||||||||||||||||||||||                      ||||||||
Db        166 TKGLNFAKETAGTNGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDKKKRAASVKDV 225
```

```
Qy          88 LNAGWNIKGVK-------NVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKIGAKT 140
               ||||||||||||       ||||| ||||||||||||||||||||||||:|||||||||
Db         226 LNAGWNIKGVKPGTTASDNVDFVHTYDTVEFLSADTKTTTVNVESKDNGKRTEVKIGAKT 285

Qy         141 SVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADK 200
               ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db         286 SVIKEKDGKLVTGKGKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTGQADK 345

Qy         201 F 201
               |
Db         346 F 346


RESULT 15
Q9JPS1
ID   Q9JPS1      PRELIMINARY;      PRT;   530 AA.
AC   Q9JPS1;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Outer membrane protein GNA992.
GN   GNA992.
OS   Neisseria meningitidis.
OC   Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC   Neisseriaceae; Neisseria.
OX   NCBI_TaxID=487;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NGF26;
RX   MEDLINE=20175756; PubMed=10710308;
RA   Pizza M., Scarlato V., Masignani V., Giuliani M.M., Arico' B.,
RA   Comanducci M., Jennings G.T., Baldi L., Bartolini E., Capecchi B.,
RA   Galeotti C.L., Luzzi E., Manetti R., Marchetti E., Mora M., Nuti S.,
RA   Ratti G., Santini L., Savino S., Scarselli M., Storni E., Zuo P.,
RA   Broeker M., Hundt E., Knapp B., Blair E., Mason T., Tettelin H.,
RA   Hood D.W., Jeffries A.C., Saunders N.J., Granoff D.M., Venter C.,
RA   Moxon E.R., Grandi G., Rappuoli R.;
RT   "Identification of Vaccine Candidates Against Serogroup B
RT   Meningococcus by Whole-Genome Sequencing.";
RL   Science 287:1816-1820(2000).
DR   EMBL; AF226380; AAF42529.1; -.
DR   InterPro; IPR005594; YadA.
DR   Pfam; PF03895; YadA; 1.
SQ   SEQUENCE   530 AA;  55190 MW;  1F836CA57598515B CRC64;

  Query Match             90.7%;  Score 923.5;  DB 2;  Length 530;
  Best Local Similarity   78.8%;  Pred. No. 3.9e-46;
  Matches  193;  Conservative    1;  Mismatches    4;  Indels   47;  Gaps    3;

Qy           4 TLKAGDNLKIKQ------------------FTYSLKKDLTDLTSVGTEKLSFSANGNKVN 45
               ||||||||||||                  ||||||||||||||| |||||| |||||||
Db          34 TLKAGDNLKIKQNTNKNTNENTNENTNDSSFTYSLKKDLTDLTSVETEKLSFGANGNKVN 93

Qy          46 ITSDTKGLNFAKETAGTNGDTTVHLNGIGSTLTD----------------------RAAS 83
               ||||||||||||||||||||||||||||||||||                      ||||
Db          94 ITSDTKGLNFAKETAGTNGDTTVHLNGIGSTLTDTLLNTGATTNVTNDNVTDDKKKRAAS 153

Qy          84 VKDVLNAGWNIKGVK-------NVDFVRTYDTVEFLSADTKTTTVNVESKDNGKKTEVKI 136
               ||||||||||||||       ||||| ||||||||||||||||||||||||:|||||
Db         154 VKDVLNAGWNIKGVKPGTTASDNVDFVHTYDTVEFLSADTKTTTVNVESKDNGKRTEVKI 213

Qy         137 GAKTSVIKEKDGKLVTGKDKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTG 196
               |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db         214 GAKTSVIKEKDGKLVTGKGKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTTANGQTG 273

Qy         197 QADKF 201
               |||||
Db         274 QADKF 278


Search completed: October  6, 2003, 09:30:55
Job time : 15.7627 secs
```

THIS PAGE BLANK (USPTO)